PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/63, 15/67, 15/70, 15/72, 15/12, 1/21, C07K 14/805 // (C12N 1/21, C12R 1:19)

A1

(11) International Publication Number: **WO 97/04110**

(43) International Publication Date: 6 February 1997 (06.02.97)

(21) International Application Number: PCT/US96/11600

(22) International Filing Date: 12 July 1996 (12.07.96)

(30) Priority Data:
60/001,179 14 July 1995 (14.07.95) US

(60) Parent Application or Grant
(63) Related by Continuation
US 60/001,179 (CIP)
Filed on 14 July 1995 (14.07.95)

(71) Applicant *(for all designated States except US)*: SOMATOGEN, INC. [US/US]; Suite FD-1, 2545 Central Avenue, Boulder, CO 80301 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: WEICKERT, Michael, J. [US/US]; 6013 Brandywine Court, Boulder, CO 80301 (US). GLASCOCK, Christopher, B. [US/US]; 1450 Oakleaf Circle, Boulder, CO 80301 (US).

(74) Agents: BROWN, Theresa, A. et al.; Somatogen, Inc., Suite FD-1, 2545 Central Avenue, Boulder, CO 80301 (US).

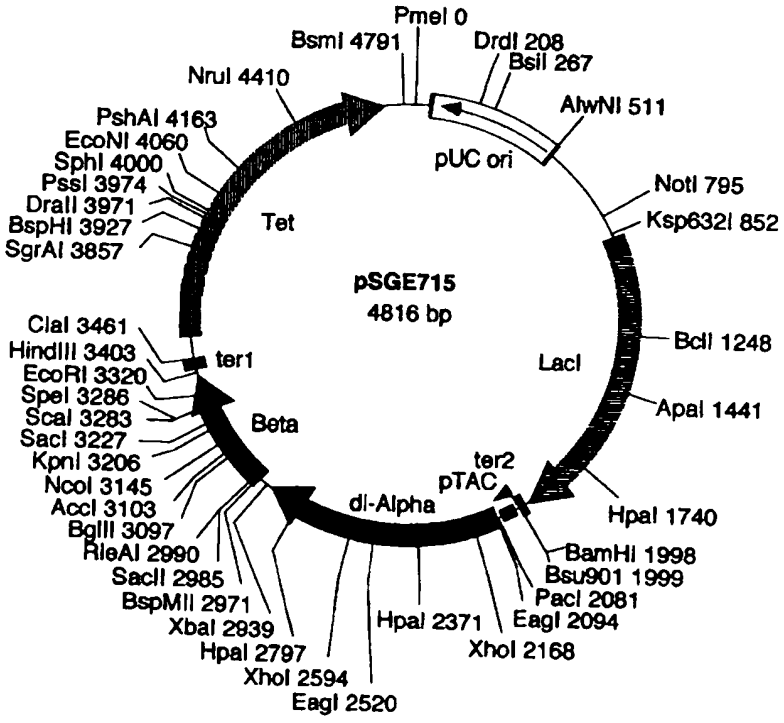
(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: METHODS FOR INCREASING PROTEIN EXPRESSION

(57) Abstract

The present invention relates to host cells and methods for expressing heterologous polypeptides. The host cell comprises a high copy number plasmid encoding said polypeptide and a chromosomally-located regulator of expression of said polypeptide under control of a strong promoter.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

METHODS FOR INCREASING PROTEIN EXPRESSION

Field of the Invention

This invention relates to expression of heterologous polypeptides. More specifically, the invention discloses novel recombinant host strains comprising high copy number cloning plasmids comprising a gene encoding the heterologous polypeptide. Expression of the heterologous polypeptide is effectively controlled by locating the regulator gene, which controls expression of the heterologous polypeptide, on the host chromosome and under control of a strong promoter.

Background of the Invention

Methods for recombinantly-producing heterologous polypeptides are well-known (see, e.g., Watson et al., Recombinant DNA, 2nd ed., (1992)). The most commonly used regulated heterologous gene expression systems in, for example, *Escherichia coli* host cells are those based on LacI-mediated repression of a promoter which controls expression of the heterologous polypeptide. Release from repression occurs through addition of an inducer (such as isopropylthiogalactoside (IPTG)), which binds to the repressor, or through an excess of DNA binding sites titrating the repressor.

Most commonly used host cells cannot control the expression of the genes cloned onto high copy number plasmids because the number of DNA binding sites on the plasmids exceeds the level of available repressor. If cloned products are deleterious or disadvantageous in a cell population, cells either die or mutations inactivating the heterologous gene accumulate.

The P_{tac} promoter in *E. coli* is 3-fold stronger than the P_{lac} promoter when fully derepressed. Therefore, it is frequently used for promoting high level regulated gene expression in E. *coli*. However, while the P_{lac} promoter is 1,000-fold repressed by LacI, the P_{tac} promoter is only 50-fold repressed (Lanzer, M. & H. Bujard. 1988. Proc. Natl. Acad. Sci. USA. **85**:8973). Repression of the E. *coli* P_{tac} promoter (De Boer, et al., 1983. Proc. Natl. Acad. Sci. USA. **78**:21-25) or other *lac*-related promoters, depends upon the concentration of the repressor, LacI. As set forth above, release from repression can occur through addition of inducer which reduces the affinity of the repressor for its specific DNA binding site, or through a reduction in the concentration of the repressor, relative to the molar concentration of specific DNA binding sites on the plasmid. If the LacI gene is located on a high copy number cloning plasmid, a large amount of inducer is required to initiate expression because of the large amount of repressor produced in such a system.

The control of P_{tac} on a multi-copy plasmid is often mediated by LacI expressed from a *lacI* promoter mutation on the chromosome, designated *lacI*^{Q}(sometimes referred to as lacI^{q} or variations thereof herein) , in a system without a corresponding *lacI* gene on the expression plasmid. The *lacI*^{Q} mutation is a single CG to TA change at -35 of the promoter region of *lacI* (Calos, M. 1978. Nature **274**:762) which causes a 10-fold increase in LacI expression (Müller-Hill, B., et al. 1968. Proc. Natl. Acad. Sci. USA. **59**:1259). Wild-type cells have a concentration of LacI of 10^{-8} M or about 10 molecules per cell, with 99% present as a tetramer (Fickert, R. & B. Müller-Hill 1992. J. Mol. Biol. **226**:59). Cells containing the *lacI*^{Q} mutation contain about 100 molecules per cell or 10^{-7} M LacI.

The plasmid pBR322 is present in 39-55 copies per cell, depending on the growth rate of the cells (Lin-Chao, S. & H. Bremer. 1986. Mol. Gen. Genet. **203**:143). Therefore, pBR322-based plasmids should be controlled by chromosomal *lacI*^{Q}. Plasmids based on pUC, having a higher copy number (500-2,000 copies per cell), should not be controlled by *lacI*^{Q} alone, and expression should be leaky. Other systems have been developed, such as the T7 system of Studier and Moffatt (1986. J. Mol. Biol. **189**:113-130 and Dubendorff and Studier. 1991. J. Mol. Biol. **219**:45-59), or involving *lacI*^{Q} overexpression on a plasmid (Law et al., 1993. J. Mol. Biol. **230**:161-173), to manage the expression of genes which require tight control of expression before induction.

However, in cases where the repressor gene is on a high copy number plasmid comprising the heterologous gene, the repressor is over-expressed relative to the amount needed to control expression of the heterologous gene controlled thereby (i.e., the stoichiometry is incorrect). As set forth above, this system requires a large amount of inducer (e.g., IPTG) to be added to the host cell during fermentation, which is costly and not efficient. On the other hand, when the repressor gene is located on the chromosome and not on the high copy number plasmid, the amount of repressor produced may be too little to control expression of the heterologous gene on the plasmid, resulting in leaky control of expression. For example, although *lacI*^{Q} results in a ten-fold higher level of expression of Lac repressor, this is not sufficient for tightly regulated expression from P_{tac} on a high copy number plasmid if *lacI*^{Q} is located on the chromosome.

In response to these and other problems, the present invention provides recombinant host strains for expressing heterologous polypeptides, comprising high copy number cloning plasmids comprising a gene encoding a heterologous polypeptide. Expression of the heterologous polypeptide is effectively controlled by locating the regulator gene, which controls expression of the heterologous polypeptide, on the host chromosome and under control of a strong promoter. A

specific embodiment relates to an allele of *lacI* capable of repressing expression of genes from *lac*-regulated promoters on high copy number plasmids. *lacIq1* (sometimes referred to as lacI^{Q1} or other variations herein) is a mutation of the repressor's promoter (Carlos and Miller. 1981. Mol. Gen. Genet. **183**:559-560) which results in ~170-fold more LacI than the wild type promoter and compared to a 10-fold increase reported for *lacq*. The increased expression of Lac repressor allows for repression of gene expression on the high copy number plasmid.

The specific embodiment provides the following representative advantages over the prior art:

- Ideal for all standard blue/white cloning and selection systems;
- Low level of *Plac* and related promoter expression in the absence of inducer;
- Safer for cloning deleterious gene products; reduces selective disadvantage of the cloned gene;
- Induction is less damaging to cells than the T7 expression system of Studier and Moffatt (Dubendorff and Studier. 1991. J. Mol. Biol. **219**:45-59) (Studier and Moffatt. 1986. J. Mol. Biol. **189**:113-130);
- Low inducer (IPTG) levels are required to induce gene expression;
- Expression level is titratable by IPTG concentrations for ideal expression levels; and
- *lacq1* is easily and rapidly detected by a PCR test.

Summary of the Invention

The present invention relates to methods and host cells for expressing heterologous polypeptides.

More specifically, the invention relates to recombinant host cells comprising high copy number cloning plasmids comprising a gene encoding a heterologous polypeptide.

More specifically, the invention relates to a prokaryotic host cell for producing a heterologous polypeptide comprising a high copy number plasmid comprising a regulatable expression unit encoding the polypeptide and a chromosomally-located gene encoding a regulatory protein capable of regulating the regulatable expression unit, expression of the regulatory protein being controlled by a strong promoter.

More specifically, the invention relates to a prokaryotic host cell for producing a heterologous polypeptide comprising a high copy number plasmid comprising a repressible expression unit encoding said polypeptide and a chromosomally-located gene encoding a repressor capable of repressing said

repressible expession unit, expression of said repressor being controlled by a strong promoter.

More specifically, the host cell is a bacterium, preferably *E. coli*, expression of the heterologous polypeptide is under control of a *LacI*-repressed promoter, and expression of the *LacI* repressor is under control of the *LacIQ1* promoter.

The invention also relates to methods for expressing heterologous proteins comprising culturing the above-described host cells and inducing expression of said genes.

The present invention also relates to methods for controlling plasmid expression. More particularly, the methods involve the use of a repressor together with an inducer to control the high-copy-number-plasmid expression of proteins, preferably recombinant proteins.

The methods can be used in any organism for which regulated promoter systems exist for heterologous gene expression, including, for example, species of the genera *Escherichia, Salmonella, Bacillus, Clostridium, Streptomyces, Staphylococcus, Neisseria, Lactobacillus, Shigella,* and *Mycoplasma.*

Brief Description of the Figures

Figure 1 is the plasmid map for pSGE705.

Figure 2 shows a simplified schematic for the construction of expression vector, pSGE705. Figures 2A shows the ligation of pBR322 origin and Tet resistance gene (PCR fragment). Figure 2B shows the insert LacI from pRGT (PCR fragment). Figure 2C shows the insert di-alpha and beta with new promoter region and shorter intergenic spacer. Figure 2D shows the results of site-directed mutagenesis to optimize ribosome binding sites and several other modifications.

Figure 3 shows a simplified schematic for the construction of expression vector, pSGE715. Figure 3A shows the ligation of pUC19 origin and Tet resistance gene (PCR fragment). Figure 3B shows the insert LacI from pRG1 (PCR fragment). Figure 3C shows the insert Ptac, di-alpha and beta from pSGE705 into BamHI and HindIII sites of pSGE508.

Detailed Description of the Invention

As used herein, the term "regulatable expression unit" means a nucleic acid molecule comprising a unit of gene expression and regulation, including heterologous genes, regulator genes and control elements which are necessary for transcription, translation and for recognition by regulator gene products, including

repressors. All of the control elements may be contiguous to the heterologous gene or not.

The term "heterologous," when referring to a gene, indicates that the gene has been inserted into a host cell either by way of a stable plasmid or through integration into the genome, and, when referring to a protein or polypeptide, indicates that the protein or polypeptide is the product of a heterologous gene.

Heterologous proteins are proteins that are normally not produced by a particular host cell.

"Host cell" refers to a prokaryotic organism containing, or hosting, a plasmid artificially constructed using techniques well known in molecular biology. Examples include *Escherichia, Salmonella, Bacillus, Clostridium, Streptomyces, Staphyloccus, Neisseria, Lactobacillus, Shigella,* and *Mycoplasma. E. coli* strains include BL21(DE3), C600, DH5αF′, HB101, JM83, JM101, JM103, JM105, JM107, JM109, JM110, MC1061, MC4100, MM294, NM522, NM554, TG1, χ^{1776}, XL1-Blue, and Y1089^{+}, all of which are commercially available, for example, from New England Biolabs (Internet address: http://www.neb.com).

"High copy number plasmid" refers to a plasmid for use in cloning heterologous genes in a host cell that is present in greater than about 200 copies per cell. Examples include pALTER®-1 Vector, pALTER®-Ex1 Vector, pALTER®-Ex2 Vector, pGEM®-3Z vector, pGEM®-3Zf(+/-) Vectors, pGEM®-7Zf(+/-) Vectors, pGEM®-9Zf(-) Vector, pGEM®-11Zf(+/-) Vectors, and the other pGEM® vectors available from Promega, the pUC-based cloning vectors, including pUC 8, 9, 18 and 19, pBacPAK8/9, and pAcUW31, and pET all available from Clontech.

"Medium copy number plasmid" refers to a plasmid for use in cloning heterologous genes in a host cell that is present in about 25-200 copies per cell. Examples include pBR322, pMAM, pMAMneo, pEUK-C1, pPUR, pYEUra3, pDR2*, pKK233-2, and pKK388-1 (all available from Clontech).

"Strong promoter" refers to a promoter or promoter mutation that increases promoter activity, i.e. transcription initiation, at least 10-fold greater than the native promoter or equivalent, or stronger than the *E. coli* lac promoter, fully induced. When referring to the promoter which controls expression of a chromosomally-located repressor, the strong promoter is called a "strong repressor promter." Examples include the lac, tac, trc, trp, ara, fru, gal, and mal promoters.

"Regulator" or "regulatory protein" refers to a transcriptional regulatory protein which exerts direct control over gene expression by binding and/or releasing DNA at a specific promoter. In particular in the context of the invention, the regulator is involved in controlling expression of the heterologous gene.

"Regulatable gene or expression unit" refers to genes or expression units effected by a regulator.

"Repressor" refers to a DNA regulatory protein which exerts direct negative control over gene expression at a specific promoter. In particular in the context of the invention, the repressor controls expression of the heterologous gene.

"Repressible" refers to a regulatory system in which the product of a regulator gene (the repressor) blocks transcription of a particular gene, or expression unit.

"Inducer" refers to an effector or molecule which inactivates the repressor and thereby causes expression of the repressible gene.

Recombinant prokaryotic systems particularly bacterial systems for producing heterologous proteins or polypeptides are well known in the art. The genes encoding the target protein can be placed in a suitable expression vector or plasmid and inserted into a microorganism or host cell. These host cells may be produced using standard recombinant DNA techniques and may be grown in cell culture or in fermentations. For example, human alpha and beta globin genes of human hemoglobin have been cloned and sequenced by Liebhaber et al. (Proc. Natl. Acad. Sci. USA 77:7054-7058, 1980) and Marotta et al. (J. Biol. Chem. 252:5040-5053, 1977) respectively. Techniques for expression of both native and mutant alpha and beta globins and their assembly into hemoglobin are set forth in U.S. Patent 5,028,588 to S.J. Hoffman, incorporated herein by reference; K. Nagai and Hoffman, S.J. et al., PCT/US90/02654; Townes, T.M. and McCune, S.L., PCT/US91/09624; and De Angelo, J. et al., PCT/US91/02568 and PCT/US91/08108 also incorporated herein by reference. The host cells of the present invention can be cultured at a temperature of about, 20°C to about 30°C, more specifically about 24°C to about 28°C and preferably about 26°C.

The present invention relates to regulation of the expression of such target heterologous proteins. One method for increasing the amount of heterologous protein expressed by a host cell involves using a high copy number plasmid into which is cloned the gene encoding the heterologous protein. Unfortunately, however, if expression of the heterologous protein is not tightly controlled, then early expression of the protein can have deleterious effects on the host cells.

To overcome these problems, a system has been invented to tightly control expression of the heterologous protein from the high copy number plasmid. The system employs a chromosomally located gene encoding a regulator of expression of the heterologous protein which is under control of a strong promoter.

The system is useful in a wide variety of prokaryotic host cells as set forth above.

High copy number cloning plasmids useful in these hosts also are set forth above.

Strong promoters and associated repressors useful together with the plasmids are set forth, for example, in Weickert & Adhya, (Weickert & Adhya. 1992. J. Biol. Chem.. 267:15869-74) incorporated herein by reference. In general, any regulator,

either a repressor or activator, with trans acting regulatory circuit can be used in the present methods and host cells.

Various inducers can be used, including various sugars such as galactose, arabinose, maltose, raffinose, ribitol, sucrose and the like, as well as purines, nucleosides, hypoxanthine, guanine, modified amino acids referred to as opines, and heavy metals. These and other useful inducers are described in Weickert & Adhya, supra.

In general, the system involves; (i) any regulatory protein which can be used to modulate the level of expression of a heterologous gene by binding, in a reversible manner, to a DNA binding site; (ii) sufficient quantity of this regulatory protein synthesized from a chromosomal gene(s) to match or exceed the number of DNA binding sites for said regulatory protein to which it can bind in the cell; (iii) promoter modifications to the promoter of the gene encoding said regulatory protein to insure production of a desired quantity of the regulatory protein; (iv) one or more DNA binding sites for said regulatory protein present on medium to high copy number plasmids (50-2,000 copies per cell), and; (v) one or more DNA binding sites for said regulatory protein near or in the nucleotide sequence encoding the heterologous gene(s), as part of the transcriptional control mechanism of the heterologous gene(s). The result is an inducible expression system with many different promoter/regulator combinations, in many different prokaryotic organisms.

Promoters of regulatory proteins in prokaryotes are typically too weak to allow production of sufficient amounts of regulatory protein to be equal to or in excess of the DNA sites found on medium to high copy number plasmids. For example, there are only about 10 copies of the Lac Repressor in each *E. coli* cell, far less than sufficient to repress a plasmid of 50-2,000 copies per cell containing one DNA binding site for Lac Repressor per plasmid. Thus there would be no significant regulation of Lac Repressor-regulatable genes on these plasmids in these cells.

The level of production of Lac Repressor and other regulatory proteins is typically low because the promoters directing transcription of the genes encoding the regulatory proteins are weak (Table A). The DNA sequences of typical promoters for expression of regulatory proteins are poor matches to the sequences efficiently recognized by RNA polymerase (Table A). Mutations which modify the promoter can produce a DNA sequence which is more efficient at elevating the transcription of the gene encoding a regulatory protein. Examples of these mutations include $lacI^Q$ and $LacI^{Q1}$ (Table A). These mutations improve the match of the -35 region of the promoter with the consensus sequence (Table A). The -35

regions of the other promoters shown, and most other regulatory protein promoters, can be similarly improved by mutations which improve the match between the promoter regions and the consensus sequences at the -35 region and also the -10 region. Other mutations can also affect the promoter activity, by, for example, changing the spacing between the -10 and -35 sequences, and changing the nucleotide sequence around the -10 region and the transcription start site. Although a particularly useful way to improve transcription and therefore production of a regulatory protein may be to engineer, by well known recombinant DNA techniques such as site directed mutagenesis, a consensus -35 region, other mutations in the entire promoter region also may provide promoter improvements. These promoter improvements are referred to as "strong promoters" "strong regulatory promoters", or "strong repressor promoters", depending upon their functions.

Promoter regions in other prokaryotic organisms may similarly be improved to make strong promoters. In Table A, examples from *Bacillus subtilis* and from *Klebsiella pneumonia* are included. The most common (vegetative) promoters for these organisms are similar to those of *E. coli*, although the strength of the the *B. subtilis* promoter is also dependent upon the DNA sequence adjacent to the -10 region, especially the proximal 5 basepairs (the extended -10 region). The strength of the promoter for the gene encoding the regulatory protein of interest can be improved by i) improving the DNA sequence homology between the promoter consensus sequence and the regulatory protein gene's promoter sequence, and/or ii) improving the promoter strength by mutation of sequence outside the consensus sequences. Note that the consensus sequences of promoters in organisms other than *E. coli* may not necessarily match the consensus -10 and -35 sequences shown in Table A. This is especially true for organisms with a substantially different nucleotide bias, for example, thermophilic organisms which have a higher G+C content than *E. coli*.

Table A.

Comparison of promoters for regulatory proteins to the consensus sequences of promoters recognized by the most abundant (vegetative) form of RNA polymerase.

	-35	-10
Consensus:	TTGACA	TATAAT
PfruR:	AAAAGCGGTTGCCGGAGTGATCAAACTGCGCTTAGATGTTAACGATTTTAA	
PmalI:	TAAATCACATTACGCAACGATAATAGCGGGTATAAGATAAATAAAAGGTAA	
PrafR:	CTTTGGTGACGGAATTTTCTGGATTTCCGGCTTAGAACCACAGCAGGAGATA	
PlacI:	CATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAG	
PlacIQ:	CATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAG	
PlacIQ1:	CATTTACGTTGACACCACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAG	
PgalR:	GCCACCCCTTGAACCAACGGGCGTTTTCCGTAACACTGAAAGAATGTAAGCG	
PgalS:	TGACTCGATTCACGAAGTCCTGTATTCAGTGCTGACAAAATAGCCCGCCAGC	
PccpA:	AGTATACGTTTTCATCATCTATAAAAACGTGTATAATTTCATGAGAAGTAAT *B.s.*	
PlacI:	GGGCGAAGCGCTGTTTTTTGTCGCGCGTTAAACATAAAATGTTAGCGCACGA *K.p.*	

Actual transcripiton start sites are in bold.

-10 and -35 sequences are underlined

B.s. = *Bacillus subtilis*

K.p. = *Klebsiella pneumonia*

Since stoichiometry is important, too much regulatory protein needs to be avoided if possible. For example, if the strength of a promoter for a gene encoding a regulatory protein is increased too much, then the abundance of the regulatory proteins may be far greater than the number of sites to which they can bind, even on a high copy number plasmid. In this case, it may require a large amount of inducer to activate transcription of the heterologous gene(s), often an undesirable situation.

After the heterologous protein has been expressed to the desired level, it generally should be released from the cell to create a crude protein solution. This can usually be done by breaking open the cells, e.g., by sonication, homogenization, enzymatic lysis or any other cell breakage technique known in the art. The proteins can also be released from cells by dilution at a controlled rate with a hypotonic buffer so that some contamination with cellular components can be avoided (Shorr et al., US Patent 5,264,555). In addition, cells may be engineered to secrete the protein of interest by methods known in the art.

After breakage of the cells, or secretion, the target protein is contained in a crude cell lysate or a crude cell broth or solution. The protein then may be purified according to methods well known in the art. For example, useful purification methods for hemoglobin-like proteins are taught in PCT Publication WO 95/14038,

incorporated herein by reference. Briefly, the methods described therein involve an immobilized metal affinity chromatography resin charged with a divalent metal ion such as zinc, followed by anion exchange chromatography. According to this publication, the solution containing the desired Hb-containing material to be purified can first be heat treated to remove Protoporphyrin IX-containing Hb. This basic purification method can be further followed by a sizing column (S-200), then another anion exchange column. Alternatively, this solution can be separated into molecular weight fractions using ion exchange chromatography. The resulting solution can then be buffer exchanged to the desired formulation buffer.

Appropriate recombinant host cells can be produced according to conventional methods or as described in the Examples below. Any suitable host cell can be used to express heterologous polypeptides. Suitable host cells include, for example, *E. coli* cells are particularly useful for expressing the heterologous polypeptides.

The proteins, so-produced, can be used for their known purposes. For example, the hemoglobin-like proteins and compositions containing the globin-like polypeptides or the multimeric hemoglobin-like proteins (collectively "hemoglobins") can be used for *in vitro* or *in vivo* applications. Such *in vitro* applications include, for example, the delivery of oxygen for the enhancement of cell growth in cell culture by maintaining oxygen levels *in vitro* (DiSorbo and Reeves, PCT publication WO 94/22482, herein incorporated by reference). Moreover, the hemoglobin-like protein can be used to remove oxygen from solutions requiring the removal of oxygen (Bonaventura and Bonaventura, US Patent 4,343,715, incorporated herein by reference) and as reference standards for analytical assays and instrumentation (Chiang, US Patent 5,320,965, incorporated herein by reference) and other such *in vitro* applications known to those of skill in the art.

The hemolgobin-like proteins also can be formulated for use in therapeutic applications. Example formulations are described in Milne, et al., WO 95/14038 and Gerber et al., PCT/US95/10232, both herein incorporated by reference. Pharmaceutical compositions comprising hemoglobin-like proteins can be administered by, for example, subcutaneous, intravenous, or intramuscular injection, topical or oral administration, large volume parenteral solutions useful as blood substitutes, etc. Pharmaceutical compositions can be administered by any conventional means such as by oral or aerosol administration, by transdermal or mucus membrane adsorption, or by injection.

For example, hemoglobins can be used in compositions useful as substitutes for red blood cells in any application where red blood cells are used, or for any application in which oxygen delivery is desired. Such hemoglobins, formulated as

red blood cell substitutes, can be used for the treatment of hemorrhages, traumas and surgeries where blood volume is lost and either fluid volume or oxygen carrying capacity or both must be replaced.

A restriction map of the entire *E. coli* genome was determined from a set of ordered Lambda clones by Kohara et al, (1987). DNA sequences of known genes have been placed on this ordered map (Rudd et al, 1992. Alignment of *E. coli* DNA sequences to a revised, integrated genomic restriction map, p. 2.3-2.43. In J. Miller (ed.), A short Course in Bacterial Genetics: a Laboratory Manual and Handbook for *Escherichia coli* and Related Bacteria. Cold Spring Harbor Prss, Cold Spring Harbor, NY). The combination of known restriction site sequence with known gene sequences was used to create a variation on anchored PCR, which was then used to clone the DNA upstream of the *lac* operon in *E. coli*. Anchored PCR relies on two primers, one of known sequence and one of random sequence. Because the sequence of the flanking restriction sites is known in *E. coli*, a portion of the sequence of the second, random PCR primer could be specified. The size of the expected fragment could be predicted, depending upon the position of the second primer designed from the known DNA sequence.

The sequence upstream of the *lacI* gene in *E. coli* was determined because of its potential role regulating the amount of Lac repressor synthesized within the cell. The sequence of the *lacI* gene is known (Calos, 1978; Farabaugh, 1978), but the published sequence ends just beyond the *lacI* promoter region. As part of the process of making the expression system, a region from a wild type *E. coli* was cloned and sequenced to establish sequence from which primers could be designed that would allow PCR amplification of *lacI* alleles from several strains. An anchored PCR technique was developed in which one primer contained a known sequence internal to *lacI* and primed toward the promoter region, and a second primer was partially random and primed upstream of the promoter region, towards the promoter. The second primer, 12 nucleotides long, contained a fixed half and a random half. The 3' (priming) end contained 6 nucleotides from a restriction site known to be upstream of P*lacI*, and the 5' 6 nucleotides were composed of equal proportions of each nucleotide at each of the six positions. A product was obtained and re-amplified by PCR using a second primer homologous to a different *lacI* sequence from the first primer, but also containing a second site for subcloning into pUC19. The insert was sequenced and contained the expected *lacI* sequence and 675 bp upstream of the *lacI* promoter region, which was not homologous to any sequence in the nucleotide databases. A large open reading frame, divergent from *lacI* , was also observed. Using this sequence, a specific nucleotide primer was designed, synthesized, and used to amplify P*lacI* from the *E. coli* chromosome of

several strains. PCR fragments of the expected size were observed from the wild type control strains, and a fragment polymorphism was used to screen for a 15 basepair P*lacI* deletion in several strains, similar to *lacI*^{Q1}.

This technique may be used to close small sequencing gaps when a map is known, and quickly extend cloning and/or sequencing into mapped regions adjacent to known sequence, from any DNA source.

The following examples are provided by way of describing specific embodiments of the present invention without intending to limit the scope of the invention in any way.

EXAMPLE 1

Constructions of plasmids containing mutations in copy number: medium copy number plasmid; pSGE705

Recombinant hemoglobin-like protein (rHb1.1, described in, for example, PCT publication WO 90/13645, incorporated herein by reference) was produced by fermentation of several *E. coli* strains containing modifications of the *lac* promoter region and/or location. Construction of the parent *E. coli* strain 1661 carrying the plasmid pSGE705 is described below. On January 20, 1994 *E. coli* strain SGE1661 was deposited with the American Type Culture Collection (ATCC Accession Number 55545). Note that Strain SGE1661 carrying the plasmid pSGE705 was denoted SGE1662 (described in PCT publication number WO 95/14038, incorporated herein by reference). pSGE705 was a medium copy number plasmid because it is present in approximately 100 copies per cell. The plasmids used in preparing pSGE705 are identified in Table 1, which also provides a brief description of each.

Materials. pBR322, pUC19 and pNEB193 were purchased from New England Biolabs, Beverly, Massachusetts. Oligonucleotides were synthesized on an Applied Biosystems DNA Synthesizer Model 392. The oligonucleotides used in preparing pSGE705 are listed in Table 2. Restriction endonucleases were purchased from New England Biolabs (Beverly, Massachusetts) and used according to manufacturer's directions. T4 DNA Ligase was purchased from either New England Biolabs or Gibco-BRL (Gaithersburg, Massachusetts) and used according to manufacturer's directions. Pfu polymerase was purchased from Stratagene (La Jolla, California) and used according to manufacturer's directions.

Media used to culture the strains are described in J. H. Miller, *Experiments in Molecular Genetics*. (Cold Spring Harbor Press 1972). and J. H. Miller, *A Short Course in Bacterial Genetics*. (Cold Spring Harbor Press 1992). Acridine orange, ampicillin

and kanamycin sulfate were purchased from Sigma Chemical Co. (St. Louis, Missouri). Tetracycline was purchased from Aldrich Chemicals (Milwaukee, Wisconsin).

Table 1. Plasmids

PLASMID	DESCRIPTION
pSGE1.1E4	rHb1.1 expression plasmid containing di-alpha and beta globin genes and resistant to both ampicillin and tetracycline
pSGE1.1E5	like pSGE1.1E4 is ampicillin resistant only
pSGE490	pUC19 lacI on a *Bam*HI-*Hin*dIII fragment
pSGE491	pUC19 α on an *Eco*RI-*Xba*I fragment
pSGE492	pNEB193 P_{tac}-α
pSGE493	pUC19 β on an *Xba* I-*Hin*dIII fragment
pSGE500	pUC19 α β on a *Bam*HI-*Hin*dIII fragment
pSGE504	pSELECT-1 replace Sty I with a Pme I site
pSGE505	pSGE504 rrnB T1 transcriptional terminator in the *Eco* RI-*Cla*I sites
pSGE507	ColE1 *ori* and *tet*, 2213 bp
pSGE509	ColE1 *ori tet lacI*, 3425 bp
pSGE513	ColE1 *ori tet lacI* α β, 4386 bp
pSGE515	ColE1 *ori tet lacI di*α β , 4812 bp
pSGE700	pTZ18U + *di*α β from pSGE515
pSGE705	modified rHb1.1 expression plasmid, ColE1 *ori* , *tet*, *lacI*, di-alpha and beta genes
pTZ18U	a phagemid derivative of pUC19, for oligonucleotide directed mutagenesis
pDLII-91F	pGEM1 + α missing valine in 2nd position (Des-val)
pNEB193	Like pUC19 but has more restriction sites in the multi cloning sites
pBR322	ColE1 *ori tet amp*
pRG1	pACYC177 *lacI*Q
pMLB1034	promoterless *lacZ* gene for promoter cloning, *amp*, pBR322 *ori*
pSGE654	constructed from pSGE705 from pSGE715. pBR322 *ori*, *tet*, *di-α*, no *lacI* gene on plasmid

Table 2. Oligonucleotides

OLIGO	SEQUENCE (5'-3')	DESCRIPTION
EV18	CGGGAATACGGTCTAGATCATTAA CGGTATTTCGAAGTCAGAACG	C-term of α gene, *Xba*I site
EV27	GATCCGAGCTGTTGACAATTAAT CATCGGCTCGTATAATGTGT GGAATTGTGACGGATAACAATTT CACACAGGAAATTAATTAATGCT GTCTCC	tac promoter *Bam*HI-sequence, *Eag*I sites
EV28	GGCCGGAGACAGCATTAATTAAT TTCCTGTGTGAAATTGTTATCCGCTCAC AATTCCACACATTATACGAGCCGATGA TTAATTGTCAACAGCTCG	tac promoter sequence, Bam HI-*Eag*I sites, complement of EV27
EV29	TCGGATTCGAATTCCAAGCTGTTGG ATCCTTAGATTGAACTGTCTCCGGCCG ATAAAACCACCG	5' end of α with *Eco*RI, *Bam*HI and *Eag*I sites
EV30	CGGAAGCCCAATCTAGAGGAA ATAATATATGCACCTGACTCCG GAAGAAAAATCC	5' end of β with *Xba*I site
EV31	CCCGAAACCAAGCTTCATTAGTGA GCTAGCGCGTTAGCAACACC	3' end of the β gene with *Hind*III site
MW007	TTTAAGCTTCATTAGTGGTATT TGTGAGCTAGCGCGT	mutagenesis reverse primer, replaces last three codons of β missing in pSGE515
MW008	CAGCATTAATTAACCTCCTTA GTGAAATTGTTATCCG	mutagenesis reverse primer to optimize α ribozyme binding site (RBS)

(Table 2 continued)

MW009	GGTGCATATATTTACCTCCTT ATCTAGATCATTAACGGTATTTCG	mutagenesis reverse primer to optimize β RBS and remove second *Bgl*II site
TG14	GGTTTAAACC	*Pme*I linker
TG59	GGCGAATAAAAGCTTGCGGCCGCG TTGACACCATCGAATGGCGCAAAA CCTTTCGCGG-	Upstream of *lacI* gene, has a *Hind*III and a *Not*I site upstream of the promoter
TG60	GGGCAAATAGGATCCAAAAAAAAG CCCGCTCATTAGGCGGGCTTTAT CACTGCCCGCTTTCCAGTCGGG	Downstream side of lacI gene with trp transcriptional terminator and a *Bam*HI site
TG62	CCCCGAAAAGGATCCAAGTA GCCGGCGGCCGCGTTCCACTG AGCGTCAGACCCC	upstream primer for pBR322 *ori* positions 3170-3148 with *Bam*HI and a *Not*I site
TG63	GGCGGTCCTGTTTAAACGCT GCGCTCGGTCGTTCGGCTGCGG	downstream primer for pBR322 *ori* positions 2380-2404 with a *Pme*I site

Genetic and Molecular Biological Procedures. Standard bacterial genetic procedures are described in J. H. Miller, *Experiments in Molecular Genetics*, (Cold Spring Harbor Press 1972) and J. H. Miller, *A Short Course in Bacterial Genetics* (Cold Spring Harbor Press, 1992). Standard molecular biology procedures were performed as described in Sambrook et al., *Molecular Cloning*, (Cold Spring Harbor Press, 1989).

Plasmid DNA Transformation. DNA transformations were performed by the procedure described in Wensick et al., Cell 3: 315-325 (1974). Briefly, cells were grown to mid log phase and then pelleted, resuspended in an equal volume of 10 mM $MgSO_4$ and incubated on ice for 30 minutes. The cells were centrifuged and the pellet resuspended in 1/2 the original volume of 50 mM $CaCl_2$ and placed on ice for

20 minutes. The cells were centrifuged again and then resuspended in 1/10 the original volume of 50 mM $CaCl_2$. Plasmid DNA was added to the competent cells in a solution of 10 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$ and 10 mM $CaCl_2$. The mixture was incubated on ice for 15 minutes and then incubated at 37°C for 5 minutes. One milliliter of LB medium was added and the mixture incubated with shaking for 30-60 minutes. The culture was then centrifuged, resuspended in 0.1 ml of LB medium and plated on the appropriate selective medium.

Purification of DNA. DNA fragments were purified from an agarose gel using the Geneclean system (Bio 101, Inc., La Jolla, CA) according to the method provided with product. PCR products were prepared and cleaved with restriction endonucleases using the Double Geneclean system. (Bio 101, Inc., La Jolla; method provided with product.) Briefly, the PCR product was purified away from the PCR primers, then the PCR product was cleaved with restriction endonuclease(s) and purified from the restriction endonuclease and buffer. The PCR product was then ready for a ligation reaction.

Annealing of oligonucleotides. Complementary oligonucleotides were annealed according to the following procedure. Equimolar amounts of each oligonucleotide were mixed in 15-25 µl of 10 mM Tris-HCl, pH 8.0/1 mM EDTA and incubated at 65°C for 30 minutes. The sample was transferred to a 37°C water bath for 30 minutes. Finally, the sample was incubated on ice for 60 minutes or in the refrigerator overnight.

Oligonucleotide directed mutagenesis. Oligonucleotide directed mutagenesis was performed with the Muta-gene phagemid *in vitro* mutagenesis kit (Bio-Rad, Hercules, California) according to the manufacturer's instructions which are based on the method of Kunkel (Kunkel, T. A. (1985) Proc. Natl. Acad. Sci. USA 82: 488; Kunkel et al., (1987) Methods Enzymol. 154: 367). The rHb1.1 region of pSGE515 was cloned into pTZ18U (Bio-Rad, Hercules, CA or U.S. Biochemical, Cleveland, OH) on a *Bam*HI-*Hind*III fragment to create pSGE700. Three oligonucleotides, MW007, MW008 and MW009 were used to simultaneously introduce multiple changes in a single reaction.

Preparation of pBR322 *ori*. PCR primers were designed to amplify the pBR322 origin of replication. These primers, TG62 and TG63, annealed to the positions 2380-2404 and 3170-3148 on the pBR322 DNA sequence (Sutcliffe, J. G.

(1979) Cold Spring Harbor Symp. Quant. Biol. 43: 77-90). The PCR product was digested with *Not*I and *Pme*I. The DNA fragment was purified according to the Geneclean procedure.

Preparation of *tet* gene fragment. The source for the *tet* gene was pSELECT-1 (Promega Corp., Madison, WI). This plasmid has a number of restriction endonuclease sites, such as *Bam*HI, *Hind*III, *Sal*I and *Sph*I removed from the *tet* gene (Lewis and Thompson (1993) Nucleic Acids Res. 18:3439-3443). A *Pme*I linker was inserted into the *Sty*I site of pSELECT-1. This plasmid was designated pSGE504. Oligonucleotides TG71: AAT TCG CGG CCG CAT TCT CGA GCG GAT CCC TGC AGC CAA GCT TAA ATA AAA CGA AAG GCT CAG TCG AAA GAC TGG GCC TTT CGT TTT ATC TGT TGA T and TG72: CGA TCA ACA GAT AAA ACG AAA GGC CCA GTC TTT CGA CTG AGC CTT TCG TTT TAT TTA AGC TTG GCT GCA GGG ATC CGC TCG AGA ATG CGG CCG CG were annealed and ligated to the *Eco*RI - *Cla*I fragment of pSGE504. The resulting plasmid, pSGE505, was shown to have the expected restriction endonuclease sites and to have lost the sites present in the multicloning site of pSELECT-1. pSGE505 was digested with *Not*I and *Pme*I. The 1417 bp fragment was purified according to the Geneclean protocol.

Preparation of lacI gene. The *lacI* gene was isolated by amplifying the gene sequence from pRG1 (Dana-Farber Cancer Inst., Boston) that carried the lacI gene. The PCR primers, TG59 and TG60 were designed to generate a wild type *lacI* promoter (Farabaugh, P. J. (1978) Nature 274:765), upstream of the gene and to place the *trp* terminator sequence (Christie et al., (1981) Proc. Natl. Acad. Sci. USA 78:4180-4184) downstream of the gene. The same step could be carried out using Y1089 (Promega) or chromosomal DNA from any *E. coli* strain carrying the *lac* region, such as MM294 (ATCC 33625.) The PCR product was gel purified and isolated according to the Geneclean procedure and cloned into *Bam*HI-*Hind*III digested pUC19 DNA to make pSGE490.

Construction of pSGE515. PCR primers EV29 and EV18 were chosen to amplify the alpha globin gene from pDLII-91F (Hoffman et al., WO 90/13645). The purified PCR product was cleaved with the restriction endonucleases *Eag*I and *Xba*I.

To create a plasmid that contained P_{tac}-α, the alpha globin gene (from above) and the tac promoter, which was prepared by annealing EV27 and EV28, were mixed with Eco RI-Xba I cleaved pUC19 DNA. The mixture of the three DNA

fragments, in approximately equimolar ratio, was treated with T4 DNA Ligase. After incubation the ligation mixture was used to transform SGE476 and ampicillin resistant transformants were selected. (Transformation into Strain MM294 (ATCC 33625) yields equivalent results.) An isolate with the correct restriction endonuclease fragments (consistent with Figure 1) was designated pSGE492. The α gene and the tac promoter DNA sequences were verified by DNA sequencing.

Primers EV30 and EV31 were used to amplify the β globin gene from pSGE1.1E4 by PCR. The purified β gene fragment was digested with *Xba*I and *Hind*III and then mixed with *Xba*I-*Hind*III digested pUC19 DNA and treated with T4 DNA ligase. The ligation mixture was used to transform competent SGE476 (equivalent to MM294, ATCC 33625) and transformants were selected on LB + ampicillin (100 μg/ml) plates. An isolate that contained the appropriate restriction endonuclease fragments (consistent with Figure 1) was chosen and designated pSGE493. The β gene was confirmed by DNA sequencing.

The β gene was isolated from pSGE493 by restriction with *Xba*I and *Hind*III followed by purification according to the Geneclean method. This DNA fragment was then ligated to *Xba*I-*Hind*III restricted pSGE492 DNA and transformed into SGE713. (Any *dam*⁻ strain such as JM110 (ATCC 47013) or GM119 (ATCC 53339) could also be used.) An ampicillin resistant transformant that carried a plasmid that had the appropriate restriction fragments (consistent with Figure 1) was chosen and designated pUC19αβ (pSGE500).

The *Bam*HI-*Hind* III fragment that contained the α and β globin genes of pSGE500 was purified according to the Geneclean method. An *Xho*I fragment that carried a portion of the di-α gene containing the glycine linker region was gel purified from pSGE1.1E5. pSGE1.1E5 (described in Hoffman et al., United States Serial Number 789,179, filed November 8, 1991) is a tetracycline sensitive analogue of pSGE1.1E4 (Hoffman et al., WO 90/13645), which could also have been used.

The pBR322 origin of replication region (pBR322 *ori*, above) was ligated to the tet gene fragment (above) and the ligation mixture was transformed into SGE476. (Transformation into MM294, above, would yield equivalent results.) Tetracycline resistant transformants were selected and plasmid DNA was isolated and analyzed. An isolate that contained the appropriate restriction endonuclease fragments was chosen and designated pSGE507.

Next, pSGE507 and SGE490 were digested with *Bam*HI and *Not*I and the appropriate fragments were purified. The two purified fragments were ligated together and the ligation mixture was used to transform competent SGE713. (Any *dam*⁻ strain could also be used; see above.) Tetracycline resistant transformants were

selected, and plasmid DNA was isolated and analyzed. A plasmid that had the appropriate restriction fragments was chosen and designated pSGE509.

The purified *Bam*HI-*Hind*III fragment of pSGE500 that contained the α and β globin genes was ligated to *Bam*HI-*Hind*III digested pSGE509. The ligation mixture was used to transform pSGE713 (see above for equivalent strains) and tetracycline resistant transformants were selected and characterized. An isolate yielding the correct size plasmid with the expected restriction endonuclease fragments was chosen and designated pSGE513.

The *Xho*I fragment of pSGE1.1E5 (described in Hoffman et al., United States Serial Number 07/789,179, filed November 8, 1991, incorporated herein by reference) that contained the di-α glycine linker sequence was ligated to *Xho*I digested pSGE513 to create a plasmid that contained the di-α gene. SGE753 was transformed with the ligation mixture and tetracycline resistant transformants were selected. (Transformation into SGE800 would have yielded equivalent results.) Isolates were screened to identify those that contained the *Xho*I fragment inserted into pSGE513 in the correct orientation. An isolate that contained the correct configuration of the di-α gene, as determined by restriction endonuclease analysis with *Eag*I, was designated pSGE515.

Modification of pSGE515 to create pSGE705. The DNA sequence record used to design PCR primers for the amplification of the β gene did not contain the C-terminal three amino acids. Oligonucleotide directed mutagenesis was used to add these nine nucleotides to the DNA sequence of the β gene. In the same reactions, modifications were introduced to optimize the ribosome binding sites for the di-α and β genes, and to remove a *Bgl*II site near the end of the di-α gene. The *Hind*III-*Bam*HI fragment from pSGE515 was subcloned into pTZ18U, creating pSGE700. pSGE700 was then used as a source of ssDNA for site-directed mutagenesis.

The following are the changes that were made with the oligonucleotides MW008 and MW009 to optimize ribosomal binding sites and to remove a *Bgl*I restriction endonuclease site.

di alpha

```
before - CAATTTCAC--AGGAAATTAATTAATGCTG
         |||||||||**||||**||||||||||||||
after  - CAATTTCACTAAGGAGGTTAATTAATGCTG
```

Four nucleotide changes, shown above, including the insertion of two nucleotides, were introduced with MW008 to optimize the ribosome binding site for di-alpha. (| - indicates identity, * - indicates a change)

beta

```
before - TAAaGATCTAGA---GGAAATAA-TATATGCAC
         |||*|||||||||***|||**|||*|||||||||
after  - TAATGATCTAGATAAGGAGGTAAATATATGCAC
```

The six nucleotide changes shown above, including the insertion of four nucleotides, were introduced with MW009 to optimize the ribosome binding site for beta. The lower case "a" on the before strand was a T to A mutation in the construction of the alpha gene that introduced a *Bgl* II site into the sequence. This was removed so that there would only be a single *Bgl* II site in pSGE705. (|)- indicates identity, * - indicates a change)

End of Beta

```
before - CTCGCTCAC---------TAATGAA
         |||||||||*********|||||||
after  - CTCGCTCACAAATACCACTAATGAA
```

MW007 introduced the coding sequence for the last three amino acids of the beta gene as shown above. (| - indicates identity, * - indicates a change)

Putative mutants were screened for loss of a *Bgl*II restriction endonuclease cleavage site (introduced by MW008). Seventeen of 24 had lost the site and were further characterized by DNA sequencing at the other two mutagenized sites. One of the 17 had incorporated all the modifications from the three oligonucleotides. These changes were verified by DNA sequencing and the rHb1.1 genes were cloned into *Bam*HI-*Hin*dIII digested pSGE509. An isolate that had the correct restriction endonuclease fragments was designated pSGE705.

A plasmid map of pSGE705 is shown in Figure 1. The plasmid map indicates many of the restriction endonuclease cleavage sites. pSGE705 is smaller than its counterpart pSGE1.1E4, and the placement of its restriction sites facilitates modular alterations of the sequence. An unused antibiotic resistance marker was removed, and a promoter was added to the *lacI* gene that would allow tighter control of rHb1.1 expression.

A new sequence upstream of the α gene minimized the distance between the tac promoter (De Boer *et al.*, Proc. Natl. Acad. Sci. USA 80, 21-25, 1983) and the first codon of the alpha gene. The intergenic region between the di-α gene and the β gene was also designed to contain the minimum sequence that contained a restriction endonuclease site and the ribosome binding site for the β gene.

EXAMPLE 2

Construction of plasmids containing mutations in copy number: High Copy Number Plasmid, pSGE720

The construction of pSGE720 was performed in two stages. First, the pUC origin of replication was introduced into PSGE705 to create plasmid pSGE715, which is similar to pSGE705 in that it includes the *lacI* gene. Then, the *lacI* gene was deleted from the plasmid and replaced with a short oligonucleotide containing several convenient restriction sites to create plasmid pSGE720.

A. Construction of pSGE715

The pUC origin of replication was introduced to create plasmid pSGE715 through pSGE508, which is identical to pSGE509 with the exception of a single basepair substitution at base 602 (G→A). The substitution changes the pBR322 origin of replication to a pUC19 origin of replication.

Plasmids pSGE508 and pSGE705 were digested to completion with restriction enzymes *Bam*HI and *Hin*dIII, according to the manufacturer's instructions (New England Biolabs.). The plasmid, pSGE508, was digested first with *Bam*HI to completion, then *Hin*dIII was added, and the digestion continued. The pSGE705 digest was purified with Promega Magic DNA Clean-up protocols and reagents (Promega, Madison, WI) and further digested to completion with *BglI*, according to the manufacturer's instructions (New England Biolabs). The enzymes in both the pSGE508 and pSGE705 digests were inactivated by heating at 67°C for 10 minutes, then the DNA was pooled and purified together using Promega Magic DNA Clean-up protocols and reagents. The DNA was suspended in ligation buffer, T4 DNA ligase was added to one aliquot, and the DNA was incubated overnight at 16°C. SGE1661 cells were made competent by the method of Hanahan, using Rubidium Chloride (Hanahan, D., In *DNA Cloning; A Practical Approach* (Glover, D. M., ed.) vol. 1, pp.109-135, IRL Press, Oxford, 1985), and transformed with the ligation mix according to the Hanahan protocol. Transformants were selected by plating the cells on LB plates containing 15μg/ml tetracycline. Candidates were screened by restriction digestion to determine the presence of the rHb1.1 genes, and sequencing to detect the pUC origin of replication. Several candidates were identified, and the resulting plasmid was named pSGE715. pSGE715 in SGE1661 was called SGE1453.

The copy number of pSGE715 is about four-fold higher than pSGE705, measured to be about 460 plasmids per cell. As noted above, the difference between pSGE705 and pSGE715 is a single basepair change in the origin of replication region, which has been confirmed by sequencing.

B. Construction of pSGE720

The *lacI* gene was deleted from pSGE715, replacing it with a short oligonucleotide containing several convenient restriction sites, by the following steps. First, plasmid pSGE715 was digested to completion with restriction enzymes *Bam*HI and *Not*I, according to the manufacturer's instructions (New England Biolabs). The pSGE715 digest was purified with Promega Magic DNA Clean-up protocols and reagents. The DNA was mixed with annealed, kinased oligonucleotides, CBG17 + CBG18, and suspended in ligation buffer.

CBG17 = 5'-GGCCGCCTTAAGTACCCGGGTTTCTGCAGAAAGCCCGCCTA
ATGAGCGGGCTTTTTTTTCCTTAGGG-3'
CBG18 = 5'-GATCCCCTAAGGAAAAAAAAAGCCCGCTCATTAGGCGGGCTTT
CTGCAGAAACCCGGGTACTTAAGGC-3'

T4 DNA ligase was added to one aliquot, and the DNA was incubated overnight at 16°C. SGE1821 cells (SGE1661 + pRG1 (pACYC177 with $lacI^Q$) (Dana-Farber Cancer Institute, Boston, MA)) were made competent by the method of Hanahan, using Rubidium Chloride, and transformed with the ligation mix according to the Hanahan protocol. SGE1821 contains pRG1 plasmids in addition to pSGE720. pRG1 is a low copy number plasmid containing *LacIq*. Transformants were selected by plating the cells on LB plates containing 15µg/ml tetracycline. Candidates were screened by restriction digestion using *Pst*I and *Sma*I to detect the presence of the new linker and the absence of the *lacI* gene, and sequenced to detect the pUC origin of replication and the absence of the *lacI* gene. Several candidates were identified, and the resulting plasmid was named pSGE720. The plasmid, pSGE720 in SGE1675 is called SGE1464.

EXAMPLE 3

Construction of Strains containing Various *lacI* Chromosomal Alleles: SGE1661, SGE1670 and SGE1675

Strains SGE1494 and SGE1495 which contained lacIq on the chromosome were purchased from ATCC, (ATCC accession numbers 47041 and 47043 respectively). SGE77 was purchased from Stratagene, Inc. (Catalogue number D1210) and also contained a *lacIq* on an F′ episome.

Strain SGE1661 contained a wild type chromosomal *lacI* allele. This strain was used as the starting material for the construction of strains containing alternative chromosomal *lacI* alleles.

To examine the effect of increasing the strength of the chromosomal *lacI* promoter, strain SGE1670 was constructed. Strain SGE1670 containing the $lacI^{q1}$ allele was constructed from SGE1661 by P1 bacteriophage transduction using a lysate grown on SGE299, selecting for kanamycin resistant transductants. Strain SGE299 contained a putative $lacI^{q1}$ allele on an F' episome adjacent a *lacZ* gene into which a *Tn5* transposon has been inserted to inactivate *lacZ*. The transposon insertion conferred kanamycin resistance to the cells. SGE299 is also know as AG1688 or RDK2759 and is described by Hu et al., (Protein Science, 2: 1072-1084, 1993). Alternatively, one skilled in the art could use any strain having a *lacI* or $lacI^{q}$ allele and mutating the sequence to yield $lacI^{q1}$ using the sequence of Calos & Miller, Mol. Gen. Genet, 183:559-560 (1981) as a guide. PCR analysis of SGE299 (described below) demonstrated that the putative $lacI^{q}$ was $lacI^{q1}$, or a variant thereof.

In strain SGE1675, the *lac* operon functionality on the chromosome was restored without affecting the *lacI* allele. The presence of the transposon insertion in *lacZ* in SGE1670 resulted in a polar mutation that destroyed the function *lacY* and *lacA* involved in lactose metabolism and transport, which may have affected the ability to induce expression of P_{tac} on the plasmid with IPTG, and thus may have affected beta-galactosidase or hemoglobin production from a given plasmid. SGE1675 was constructed by transduction of SGE1670 from a P1 lysate prepared on SGE765. SGE765 was a strain containing a *Tn5* insertion into the *lacI* gene conferring kanamycin resistance to the cells. In addition, SGE765 contained a wild type copy of *lacZ*. Strain SGE765 was made by P1 transduction from a lysate made on MS24 into strain C3000. MS24 is also known as MG1655 *lacI3098*::$Tn10^{kan}$, *acZA118*, and is described in Singer, et al. (Microbiol. Rev. 53: 1-24, 1989). Strain C3000 is available from the American Type Culture Collection, ATCC # 15597.

Transductants were selected for their ability to grow on minimal medium containing lactose as the sole carbon source and screened for sensitivity to kanamycin. Transductants that were sensitive to kanamycin and were lac^{+}, were screened for their ability to regulate (repress) the expression of beta-galactosidase from plasmid pSGE714 (see below). Note: pSGE714 does not contain a *lacI* on the plasmid and was examined for repression of *lacZ* as described for strain SGE1670.

EXAMPLE 4

Medium Copy Number LacZ Fusion Plasmid containing *lacI:* pSGE712

pSGE712 containing the *lacI* gene and a fusion of the *lacZ* gene to the 5' end of the beta globin gene was constructed as described below. This construct provided a convenient tool for screening control of expression as a function of beta-galactosidase activity in the cells. Thus in the absence of inducer, the expression of beta-galactosidase provided a sensitive measure of control of expression by the product of the *lac* repressor gene expressed primarily from the plasmid.

pSGE705 was digested with *Sal*I and *Bgl*II. pMBL1034 (obtained from Sankar Adhya, NIH, and described in Miller, G. H., Experiments in Gene Fusions, (1984), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) was digested with *Bam*HI and *Bal*I. The two digests were cleaned using a Promega MagicDNA clean-up column using the manufacturer's recommended procedures, and the DNA was ligated using T4 DNA ligase. The ligation mix was used to transform JM109 competent cells. Candidates were ampicillin resistant and blue colonies on LB amp plates containing X-gal and were screened by *Bam*HI/*Hin*dIII digestion for the appropriate plasmid conformation. Correct candidates contained a *Bam*HI-*Hin*dIII fragment almost 3 Kilobase pairs longer than the original plasmid, indicating the presence of the *lacZ* gene on the plasmid.

EXAMPLE 5

Construction of a Medium Copy Number LacZ Fusion Plasmid without *lacI:* pSGE714

pSGE714 containing a fusion of the *lacZ* gene to the 5' end of the beta globin gene was constructed as described in Example 4, except that pSGE654 was used instead of pSGE705 as the starting material. Note that unlike pSGE712, pSGE714 did not contain *lacI*. This construct provided a convenient tool for screening control of expression as a function of beta-galactosidase activity in the cells. Thus, in the absence of inducer, the expression of beta-galactosidase provided a sensitive measure of control of expression by the product of the *lac* repressor gene expressed from the chromosome of the cell.

EXAMPLE 6

Construction of a High Copy Number LacZ Fusion Plasmid without *lacI*: pSGE721

pSGE721 containing a fusion of the *lacZ* gene to the 5' end of the beta globin gene was constructed as described in Example 4, except that pSGE720 was used instead of pSGE705 as the starting material. This construct provided a convenient tool for screening control of expression as a function of beta-galactosidase activity in the cells. Thus, in the absence of inducer, the expression of beta-galactosidase provided a sensitive measure of control of expression by the product of the *lac* repressor gene expressed from the chromosome of the cell. This plasmid results in approximately 500 copies of the plasmid per cell. In order to control the expression of beta-galactosidase from these plasmid copies, more repressor must be expressed from the chromosomal copy of *lacI* per cell than there are plasmids per cell.

EXAMPLE 7

Construction of a high copy number plasmid containing a single alpha globin gene and wild type beta gene: pSGE728

The construction of pSGE728 was performed by digesting plasmid pSGE720 with an enzyme that cuts only within each of the two alpha subunits of the di-alpha gene encoding the globin-like protein, followed by ligation, to preferentially reconstruct deletions of one alpha subunit, and the di-alpha glycine linker. The resulting plasmid, pSGE726, contains a single alpha gene rather than a di-alpha gene. The Presbyterian mutation of human hemoglobin in the beta globin gene was replaced by a second digestion and ligation that introduced the wild-type beta and created pSGE728.

A. Construction of pSGE726

pSGE720 was digested with restriction enzyme *Xho*I, according to the manufacturer's instructions (New England Biolabs). The pSGE720 digest was purified with Promega Magic DNA Clean-up protocols and reagents (Promega, Madison, WI) and suspended in ligation buffer. T4 DNA ligase was added and the DNA was incubated overnight at 16°C. SGE1675 cells were made competent by the method of Hanahan (Hanahan, D, ibid), and transformed with the ligation mixture according to the referenced protocol. Transformants were selected by plating the

cells on LB plates containing 15µg/ml tetracycline. Candidates were screened by restriction digestion to determine the presence of a mono-alpha globin gene, rather than the di-alpha globin gene present in pSGE720. Several candidates were identified, and the resulting plasmid was named pSGE726 in SGE1675 was called SGE1480. This plasmid expressed the hemoglobin-like molecule rHb1.0, mono-alpha plus beta^{Presbyterian}.

B. Construction of pSGE728

pSGE726 and pSGE0.0E4 (Hoffman et al., 1990. Proc. Natl. Acad. Sci. USA. 87:8521-8525 and Looker et al., 1992. Nature, 356:258-260) were digested with restriction enzymes *Bgl*II and *Hind*III, according to the manufacturer's instructions (New England Biolabs). The digests were loaded into wells in a 1% agarose gel, and subjected to electrophoresis in Tris acetate buffer as described in Maniatis et al. (Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). The small DNA band of about 300 basepairs (bp) from the pSGE0.0E4 digest, and the large band of about 3,400 bp from the pSGE726 digest, were purified by excising them from the agarose gel, pooling them into a 1.7ml tube and purifying them with GeneClean Kit ptotocols and reagents (BIO 101, La Jolla, CA). Purified, pooled DNA fragments were suspended in ligation buffer. T4 DNA ligase was added and the DNA was incubated overnight at 16°C. SGE1675 cells were made competent by the method of Hanahan (Hanahan, D, ibid), and transformed with the ligation mixture according to the referenced ptotocol. Transformants were selected by plating the cells on LB plates containing 15g/ml tetracycline. Candidates were screened by restriction digestion with *Sca* I according to the manufacturer's instructions (New England Biolabs), to determine the presence of a wild type beta globin gene, which is not susceptible to *Sca*I digestion, rather than the beta^{Presbyterian} gene which is cleaved by *Sca*I. Several candidates were identified. The resulting plasmid was named pSGE728. pSGE728 in SGE1675 was called SGE1483. This plasmid expressed rHb0.0, mono-alpha plus wild type beta.

EXAMPLE 8

Plasmid Copy Number Measurement

Plasmid copy-number was determined by comparison of the band intensity of linearized plasmids to that of known quantities of *Hind*III-digested lambda DNA (New England Biolabs) on Ethidium Bromide-stained TAE agarose gels. Plasmid DNA was extracted from bacterial cultures using the Wizard Miniprep Kit (Promega). The OD_{600} of these cultures was determined before plasmid

purification. Using the conversion factor of 8×10^8 cells/OD*ml, the OD_{600}, the culture volume used, and the mass of DNA determined by comparison to the standard, the copy number was calculated.

EXAMPLE 9

Measurement of beta-galactosidase expression

Beta-galactosidase was measured using the technique described in Miller. Cells were grown with agitation in supplemented minimal medium (M63 + casamino acids) at 37 or 30°C in 25 ml flasks or 20 ml test tubes. Expression was induced by the addition of various concentrations of IPTG to the cells in mid-log phase (OD 600 nm >0.2, but <0.8) and incubation was continued for 3-5 hours before sampling for beta-galactosidase analysis.

EXAMPLE 10

Determination of Hemoglobin Expression

Fermentation

Hemoglobin-like proteins are expressed in the strains described above using a standard fermentation protocol. First, a fermentor inoculum is grown from seed stock. The inoculum is then transferred to a 15 liter fermentor and induced. The details of the fermentation process are described below.

Seed Stock

Seed stock is grown up in LB broth containing 10 g/L BactoTryptone™, 5 g/L yeast extract, 5 g/L NaCl, 0.2 g/L NaOH, and 10 ug/ml tetracycline to an optical density of 1.5 - 1.7 at 600 nm. The solution is then made up to 10% glycerol and stored at -80°C until required.

Fermentor Inoculum (500 ml broth in 2 L shake flasks)

To prepare the fermentor inoculum, seed stock is thawed and 0.1-0.4 ml of seed stock are inoculated into 500 ml of a solution containing approximately 4 g/L KH_2PO_4, 7 g/L K_2HPO_4, 2 g/L $(NH_4)_2SO_4$, 1 g/L Na_3 Citrate·$2H_2O$, 153 mg/L $MgSO_4 \cdot 7H_2O$, 2.3 g/L of L-proline, 2 g/L yeast extract, 4.8-5 g/L glucose, 75 mg/L thiamine HCl, 12 mg/L tetracycline, 81 mg/L $FeCl_3 \cdot 6H_2O$, 4 mg/L $ZnCl_2$, 6 mg/L $CoCl_2 \cdot 6H_2O$, 6 mg/L $Na_2MoO_4 \cdot 2H_2O$, 3.1 mg/L $CaCl_2 \cdot 2H_2O$, 3.9 mg/L $Cu(II)SO_4 \cdot 5H_2O$, 1.5 mg/L H_3BO_3, and 300 µl/L HCl. This culture is allowed to grow for 10 hours at 37°C on a shaker. Four flasks are combined and used to inoculate the 15L Fermentors.

15L Fermentor (14 L volume in 20 L Fermentor - "15L")

The entire fermentor inoculum is then asceptically transferred to a 20-liter fermentor containing 10 liters of the following: 1.8 g/L KH_2PO_4, 3.3 g/L K_2HPO_4 ,1.8 g/L $(NH_4)_2SO_4$, 155 mg/L thiamine HCl , 10.3 mg/L tetracycline, 3.1 g/L proline, 1.9 g/L $MgSO_4 \cdot 7H_2O$, 1.9 g/L Na_3-citrate$\cdot 2H_2O$, 133 mg/L $FeCl_3 \cdot 6H_2O$, 6.4 mg/L $ZnCl_2$, 9.9 mg/L $CoCl_2 \cdot 6H_2O$, 9.9 mg/L $Na_2MoO_4 \cdot 2H_2O$, 5 mg/L $CaCl_2 \cdot 2H_2O$, 6.3 mg/L Cu(II)$SO_4 \cdot 5H_2$ O, 2.5 mg/L H_3BO_3, and 494 µl/L HCl. Note that masses of added reagents are calculated using the final volume of fermentation, 14 liters, and are approximate within measurement error. The pH is maintained at 6.8 to 6.95 by addition of 15% to 30% NH_4OH, dissolved oxygen is maintained at or above 20%, and 50 to 70% glucose is added throughout the growth period, sufficient to maintain low but adequate levels of glucose in the culture (0.1 g/L-10 g/L). Dissolved oxygen is maintained as close to 20% as possible. The culture is grown between 28 and 32°C for approximately 12 hours prior to transfer to the 600 liter fermentor, if such transfer is required, or alternatively is induced by IPTG addition and grown for 10-16 additional hours as described below.

600L Fermentations

The entire seed fermentor inoculum is then asceptically transferred to a 600-liter fermentor containing approximately 375 liters of the solution containing: 1.8 g/L KH_2PO_4, 3.3 g/L K_2HPO_4, 1.8 g/L $(NH_4)_2SO_4$,3.3 ml/L polypropylene glycol-2000, 220 g/L glucose, 143 mg/L thiamine HCl, 9.4 mg/L tetracycline, 1.4 g/L $MgSO_4 \cdot 7H_2O$, 1.4 g/L Na_3-citrate$\cdot 2H_2O$, 2.9 g/L L-proline, 99 mg/L $FeCl_3 \cdot 6H_2O4$, .8 mg/L $ZnCl_2$, 7.3 mg/L $CoCl_2 \cdot 6H_2O$, 7.3 mg/L $Na_2MoO_4 \cdot 2H_2O$, 3.7 mg/L $CaCl_2 \cdot 2H_2O$, 4.7 mg/L Cu(II)$SO_4 \cdot 5H_2$ O, 1.8 mg/L H_3BO_3, and 366 µl/L HCl. Note the reagent additions are calculated with the final volume of the fermentation, 450 liters, and that all masses are approximate.

Whether at 600 L scale or at 15 L scale, the pH is maintained at 6.8 - 6.95 by addition of 15% to 30% NH_4OH, dissolved oxygen is maintained at or above 20%, and 50-70% glucose is added throughout the growth period, sufficient to maintain low but adequate levels of glucose in the culture (0.1 g/L-10 g/L). The culture is grown between 24 and 30°C to an OD_{600} ~ 10-40 prior to induction with 10-1000 µM IPTG. Upon induction of hemoglobin synthesis, the *E. coli* heme biosynthesis is supplemented by addition of hemin dissolved in 1 N NaOH, either by addition of the total mass of hemin required at induction, by continuous addition of hemin throughout the induction period, or by periodic addition of hemin dissolved in 50 mM to 1 M NaOH (e.g. one third of the total mass of hemin to be added to the fermentor is added at induction, another third is added after 1/4 of the total time

after fermentation has elapsed, and the last third is added half-way through the induction period). Total hemin added ranges from 50 to 300 mg/L. The fermentation is allowed to continue for 8-16 hours post-induction. At the end of this period, several 1 ml aliquots are removed from the broth for determination of hemoglobin production.

The fermentor is run at several temperatures (see further examples below), controlling dissolved oxygen at 20% and glucose between 0-6 g/L. At OD 30 ± 2, induction occurs by adding various amounts of 100 mM IPTG to yield several different concentrations of inducer, as further described below. In 15L fermentations at induction, 10 mL of 50 mg/mL hemin is added along with the appropriate volume of IPTG. At 3 hours post induction, 13 mL of 50 mg/mL hemin is added and at 6 hours post induction, 17 mL of 50 mg/mL hemin is added. Harvest and further purification occurs at 10-16 hours post induction.

Soluble Hemoglobin Measurement by Immobilized Metal Affinity Chromatography

Samples were collected as above and pelleted. Cells were lysed and DNA was fragmented as described below:

1. Cell Lysis

The samples were placed on ice and protected from light. 1.00 ± 0.01 mL of 25.0mM $Na_2B_4O_7 \cdot 10H_2O$ was added to the cell pellet and the sample was vortexed until the contents were resuspended. 30 ± 1 µL of Lysozyme NaCl solution was added to the resuspended pellet and the sample was re-vortexed. The samples were incubated on ice for 30 - 45 minutes while protected from light. Following this incubation, the samples were mixed at least three times and placed in a 37° ± 1°C water bath for 2.25 ± 0.25 minutes. After incubation at 37°C, the samples were again mixed.

2. DNA Fragmentation

30 ± 1 µL of DNase was added to the sample which was then vortexed. The sample was then incubated with rocking for 15 minutes at room temperature, while ensuring that the sample was protected from light.

Hemoglobin was extracted from the samples after cell lysis by first freezing at -80°C for greater than three hours. The samples were then thawed by rocking at room temperature in the dark. After thawing, samples were mixed and then centrifuged for 20 minutes at 14,000 rpm in a microfuge. The supernatant was removed and treated with CO gas for 30 seconds. This CO treated supernatant was

then heated for 4 minutes at 65°C, if necessary, then spun for 4 minutes in a microcentrifuge at 14,000 rpm. 0.1 ml of the supernatant was then diluted with 0.9 ml of 5 mM sodium phosphate, 150 mM sodium chloride, pH 7.4, and filtered into an autosampler vial.

Samples were then analyzed on a Biocad Perfusion Chromatography Workstation (PerSeptive Biosystems), equipped with a Poros (MC/M) metal chelate perfusion column, 4.6 mm x 100 mm, (PerSeptive Biosystems) charged with zinc. The column was run at a flow rate of 6 ml/min and the effluent was monitored at 412 nm. The column was loaded with 3 column volumes of 10% 2 M NaCl+ 80mM Tris-HCl, pH8.0/90% water, washed with 4 column volumes of 25% NaCl/75% water and eluted with 50% 2 M NaCl+ 80mM Tris-HCl, pH8.0/50% 50mM EDTA, pH8.5.

Soluble Hemoglobin Measurement by Difference Assay

Five microliters of hemoglobin solutions were added to 500 µl of 0.1 M Tris, pH 8.0. 200 µl of the diluted hemoglobin solution was then added to 2.8 ml of 0.1 M Tris, pH 8.0, in a 4.5 ml cuvette for a final dilution of 1:1500. The oxygenated sample (HbO2) was then analyzed by spectrophotometry in a Hewlett-Packard model HP 8452A spectrophotometer. Absorbances at 436, 425, 420, 404, 400 nm were collected and stored in a data storage system. The cuvette was then removed from the spectrophotometer and sparged with carbon monoxide two times for 15 seconds each time. The cuvette was inverted 5 times between each sparge. The sample was then re-inserted into the spectrophotometer, and a second set of spectra were collected that corresponded to carbonmonoxy hemoglobin (HbCO). The cuvette was then again removed from the spectrophotometer and 30 µl of 0.1M KCN in 0.1 M Tris, pH 8.0, was added to the sample. The sample was then inverted three times, allowed to incubate for 5 minutes, and re-inserted into the spectrophotometer for a final spectrophotometric analysis (HbCN).

The concentration of total hemoglobin was determined using the following quantities:

$$\Delta O.D.(\text{rHb}) = [A_{420}^{HbCO} - A_{420}^{HbO2}] - [A_{436}^{HbCO} - A_{436}^{HbO2}]$$

$$\Delta O.D.(\text{rHb+}) = [A_{424}^{HbCN} - A_{424}^{HbCO}] - [A_{404}^{HbCN} - A_{404}^{HbCO}]$$

where A = the absorbance at the susbcripted wavelength for the superscripted hemoglobin species.

Using these values, total reduced hemoglobin (rHb) and oxidized hemoglobin (rHb+) was calculated as follows:

$$[rHb\ (mg/ml)] = \Delta O.D.(rHb)/\ 5.132$$

$$[rHb+\ (mg/ml)] = \Delta O.D.\ (rHb+)/7.169$$

Total hemoglobin (mg/ml) is then simply [rHb] + [rHb+].

Insoluble Hemoglobin Measurement by Western Blot

Cells were lysed as described for the Immobilized Metal Affinity Chromotography method above. The soluble and insoluble fractions were separated by centrifuging the lysate for 15 minutes in a microfuge at top speed. The supernatant (soluble fraction) was transferred to another microfuge tube. The soluble fractions were assayed for total protein using the BCA (Bicinchoninic acid) assay (Pierce, #23225X, Rockford, Illinois) and diluted to 1 mg/ml with H_2O. The pellet (insoluble fraction) was resuspended with 1 ml of 1% SDS in 25 mM sodium tetraborate and diluted such that the same volume of solublized pellet and water were used as were used to dilute the corresponding soluble fraction. Each fraction was diluted 1:1 with 2x SDS sample buffer (125 mM Tris·HCl, 20% glycerol, 2% SDS, 2% beta-mercaptoethanol, 0.01% bromophenol blue). 10ml of each sample was applied to a 12% SDS-PAGE gel and run overnight (approximately 14 hours) at 80 V constant voltage. The gel was transferred onto ProBlott Membrane (Applied Biosystems #400994) in 10 mM CAPS, pH 11, 10% MeOH for two hours at 400 mA (Hoefer TE22). The blots were blocked for one hour to overnight in Tris Buffered Saline plus Tween 20 (TBST, 10 mM Tris pH 8.2, 150 mM NaCl, 0.05% Tween 20) plus 5% Food Club brand Non Fat Dry Milk (NFDM). The first antibody was an affinity purified goat anti- rHb antibody. The antibody was diluted 1:2500 in TBST plus 1% NFDM and rocked at room temperature for one hour. Prior to second antibody incubation, the blot was washed three times for five minutes per wash in TBST. The second "antibody" was Horseradish peroxidase (HRP) linked to Protein G (Calbiochem catalogue #539322, San Diego, California). Stock protein G-HRP was diluted 1:5000 with TBST plus 0.5% NFDM. Blots were incubated with rocking for 1-2 hours. Three final TBST washes of five minutes each were performed. Blots were developed for one minute with LumiGLO Substrate (Kirkegaard & Perry Laboratory, #54-61-00), covered with plastic wrap and exposed to X-Ray film (Amersham Hyperfilm-ECL #RPN. 2103).

Developed films were scanned on a pdi (Huntington Station, NY) scanning densitometer using the white light setting recommended by the manufacturer. The densitometer is pdi's Model DNA 35 connected to a Sun Microsystems SPARC

Station 2 computer running pdi's Quantity One version 2.4 software. The software generates an image of the blot which is then analyzed. Image background was subtracted along a stripe down the center of each lane, by a "rolling circle" method . Each band was quantified by average density in the band times the band area. Quantitation values are in units of O.D. x mm^2. Percent soluble recombinant hemoglobin (rHb) was calculated for each time point by the formula:

$$\% \text{ Soluble Hb} = \frac{\text{Soluble Value}}{\text{Soluble Value + Insoluble Value}} \times 100$$

EXAMPLE 11

Control of beta-galatosidase production on medium and high copy number plasmids by chromosomal lacI alleles

Production of beta-galactosidase in strain SGE77 (Stratagene) containing either plasmid pSGE712(medium copy number), 714(medium copy number), or 721(high copy number) was compared to strain SGE1495 containing either plasmid pSGE712, 714, or 721, and several other strains as well, by growing in 25 ml flasks or 20 ml test tubes as described above. At induction, the cultures were divided into two equal portions and IPTG was added, to 100 μM final concentration, to one half of the culture. The other was left untreated (control). Incubation was continued for both under identical conditions. Beta-galactosidase expression in the IPTG treated fractions was monitored to ensure that expression levels were consistently inducible. Beta-galactosidase activity was also determined from a sample of the cultures that were not treated with IPTG. In the absence of inducer, beta-galactosidase expression was lowest in strains that contained a *lacI* gene on the plasmid or in any of the strains derived from strain SGE1675 which contained a chromosomal $lacI^{q1}$ allele (Table 3). Thus, sufficient Lac repressor was produced from the chromosomal *lacI* allele to control expression from P*tac* on both medium and high copy number plasmids which did not contain *lacI*.

Table 3.

Strain/Description	Plasmid	β-galactosidase activity (Miller Units)	
		no IPTG	+100uM IPTG
SGE77/*lacIQ*	pSGE712	3,069	52,482
SGE77/*lacIQ*	pSGE714	4,247	32,783
SGE77/*lacIQ*	pSGE721	24,402	25,141
SGE1495/*lacIQ*	pSGE712	1,846	36,511
SGE1495/*lacIQ*	pSGE714	3,957	33,763
SGE1495/*lacIQ*	pSGE721	9,659	39,293
SGE1661/*lacI+*	pSGE721	20,881	31,118
SGE1674/*lacIQ1*	pSGE721	1,203	39,770
SGE1675/*lacIQ1*	pSGE721	1,778	27,124

lacI^{Q} on the chromosome did not control expression of a beta-LacZ fusion from high copy number (~500 copies per cell) plasmid (pSGE721). *lacI*^{Q} on the chromosome was able to control expression of a medium copy plasmid (~100 copies per cell; pSGE714) fairly well. A medium copy number plasmid which contained its own *lacI* gene with the native *lacI* promoter (pSGE712) was capable of controlling expression of this well (~100 copies per cell).

lacI+ or *lacI*^{Q} on the chromosome did not control expression of a beta-LacZ fusion from high copy number (~500 copies per cell) plasmid (pSGE721 in SGE77, SGE1469 and SGE1495 above). *lacI*^{Q1} on the chromosome was capable of controlling expression of the high copy plasmid the best of the strains above. Note that SGE1674 and 1675 are sibs from the same transduction.

EXAMPLE 12

Control of hemoglobin production on medium and high copy number plasmids by chromosomal lacI alleles

Production of hemoglobin in strain SGE1662 containing plasmid pSGE705 was compared to production of hemoglobin from SGE1464 containing plasmid pSGE720. Cells were grown in 15 L fermentors at 30°C as described above. Samples were withdrawn just prior to induction and assayed for hemoglobin levels on a per cell and per volume basis. There was no significant difference in hemoglobin production prior to induction between SGE1662 and SGE1464. Ten hours after induction, samples were again withdrawn and hemoglobin expression was

measured. About twice as much hemoglobin was produced in SGE1464 relative to SGE1662 (Table 4) even though SGE1464 required approximately three-fold less IPTG for induction.

Table 4.

Oxy rHb1.1 Yields by Zn Capture from 15 L Fermentations with SGE1464 and SGE1662.

SGE #	ferm. Oxy (g/L)	IPTG µM
1464	1.10±0.29	15
1662	0.58±0.17	55

This demonstrated the superiority of the high copy number plasmid expression system in which the expression of the heterologous genes, in this case those encoding recombinant hemoglobin, are controlled by a regulator overproduced from the chromosome (SGE1464). Induction of higher levels of soluble globin expression occurred even with lower concentrations of inducer than required to induce globin protein expression from the medium copy plasmid containing a lacI gene on the plasmid (SGE1662).

The level of soluble rHb1.1 produced using various low IPTG concentrations for induction was examined. Cells were grown in 15L fermentors at 30°C as described above. Ten hours after induction with IPTG concentrations ranging from 0 to 55µM, samples were withdrawn and soluble hemoglobin expression was measured. Significant induction of expression was observed with the lowest IPTG concentration used, 5.5µM (Table 5), while no hemoglobin was produced in the absence of IPTG. The hemoglobin concentration produced increased only marginally when the IPTG concentration was above 7.5µM (Table 5). The expression of rHb1.1, controlled on a high copy plasmid by Lac Repressor produced from the *lacI^{Q1}* allele on the chromosome, was sensitive to induction by low IPTG concentrations, a significant advantage over other expression systems.

Table 5.
Soluble rHb1.1 from SGE1464 induced by various IPTG concentrations.

µM IPTG	g/L rHb1.1±S.D
0	0.02±0.01
5.5	0.35
7.5	0.8
15	0.73
22.5	0.91±0.28
55	0.9

EXAMPLE 13

Effect of Temperature on soluble di-alpha hemoglobin expression in SGE1464

The effect of temperature on expression levels of soluble expression of di-alpha hemoglobin (described in co-pending PCT application WO90/13645 to Hoffman et al.) in strain SGE1464 containing pSGE720 was examined by growing the cells in 15 L fermentors at various temperatures. Soluble hemoglobin was measured using the IMAC procedure described above. Soluble hemoglobin expression was enhanced when the cultures were grown at lower temperatures. 28°C significantly improved soluble hemoglobin expression for the high copy number system (SGE1464 -Table 6).

Expression was not enhanced for the strain containing the medium copy number plasmid which also includes the Lac Repressor gene on the plasmid (Table 6; SGE1662). Because the high copy number plasmid facilitates production of more rHb1.1 due to the 4- to 5-fold increase in rHb1.1 gene dosage (on the high copy number plasmid, relative to the medium copy plasmid; 500 versus 100 copies per cell approximately), the effect of temperature on solubility of the synthesized protein is manifest. SGE1464 produces a significant excess of protein which is insoluble and transparent to the IMAC assay. We identify the insoluble material by a Western Blot assay (supra) using an antibody specific to rHb1.1. SGE1662 does not produce insoluble rHb1.1. All the globin synthesized is soluble. SGE1464 produces a substantial proportion of insoluble globin protein, usually approximately equivalent to the amount of soluble protein. Therefore, temperature, which is known to improve solubility and/or stability of some proteins, can only affect the strain with a reservior of insoluble rHb1.1, SGE1464. Overproduction of the protein through increased gene dosage and improved inducibility created an opportunity for

increased soluble globin expression that lower temperatures allowed further exploitation.

Table 6.

SGE #	ferm. Oxy (g/L)	ferm. temp.
1464	1.10±0.29	30°C
1464	1.50±0.21	28°C
1662	0.58±0.17	30°C
1662	0.51±0.11	28°C

In an additional study, the proportion of soluble rHb1.1 protein produced from the high copy number system (SGE1464) increased inversely with temperature (Table 7). At the lowest temperature tested, 24°C, an estimated 70% of the globin protein synthesized was soluble, while at 30°C, only approximately 34% was soluble (Table 7). We expect that this relationship is expected to continue if additional experiments extended the temperature extremes beyond the tested range, but the optimal yield is likely to fall within range of temperatures tested since the highest peak yield per volume of culture was obtained at 26°C (Table 7) with lower yields at either extreme of lower or higher temperatures.

Table 7.

Temp. (°C)	Ave. % sol.	Ave. peak g.L
24	70	1.10
26	62	1.62
28	53	1.28
30	34	0.83

EXAMPLE 14

Effect of Temperature on soluble hemoglobin expression in SGE1483

The effect of temperature on expression levels of soluble expression of wild-type hemoglobin containing a replacement of the N-terminal valine with methionine for all the globins (described in co-pending PCT application WO90/13645 to Hoffman et al.) in strain SGE1483 containing pSGE728 was examined by growing the cells in 15 L fermentors at various temperatures. The plasmid, pSGE728, is the same

as pSGE720 except it contains only one alpha gene (i.e., a mono-alpha gene) instead of the di-alpha gene of pSGE720. SGE1483 is the same as strain SGE1464 except SGE1483 contains pSGE728 instead of pSEG720.

Soluble hemoglobin was measured using the IMAC procedure described above. Soluble hemoglobin expression was enhanced when the cultures were grown at lower temperatures. 28°C significantly improved soluble hemoglobin expression for the high copy number system (SGE1483 -Table 8).

Table 8.

Temp. (°C)	g/L SGE1483 (rHb0.0)
30	1.60 ± 0.17
28	2.16 ± 0.29

EXAMPLE 15

Table 9.

lacI allele	plasmid	plasmids per cell	LacI per cell	LacI:Plasmid ratio
lacIQ	pSGE705	100	1100	10:1
lacIQ	pSGE720	500	100	1:5 (no control)
lacIQ	pSGE715	500	5100	10:1
lacIQ1	pSGE720	500	1700	3:1

The proportional relationship (stoichiometry) of the regulatory protein and the DNA binding site to which it binds, helps determine the level of control of gene expression in the absence of inducing conditions, and the amount of inducer required for full induction of gene expression from the promoter so-controlled. In the example system described in Table 9, LacI is the regulatory protein and the plasmid contains one DNA binding site for LacI on each copy in the cell. The plasmid copy number is estimated as described (supra), and the LacI level is estimated from relative transcription levels (infra). The ratio of LacI to plasmid must exceed 1:1 in order to control expression of gene(s) whose transcription is regulated by LacI, on every plasmid in the cell. The only example in the Table 9 which does not produce a ratio greater than 1:1 is the presence of a high copy number plasmid (pSGE720) in a strain having only *lacIQ* on the chromosome. This should result in no control of expression (i.e. the expression is induced even in the absence of IPTG

or other inducer) from at least some plasmids in the cell, since there is not enough LacI to go around. Two actual examples of this senario are shown in Table 3, in which strains containing only lacIQ (SGE77 and SGE1495) failed to repress a high copy number plasmid, pSGE721, in the absence of inducer.

Since stoichiometry is important, too much regulatory protein needs to be avoided if possible. If the strength of a promoter, or the gene dosage for a gene encoding a regulatory protein is increased too much, then the abundance of the regulatory protein may be far greater than the number of sites to which it can bind, even on a high copy number plasmid. In this case, it may require a large amount of inducer to activate transcription of the heterologous gene(s), often an undesirable situation. A large excess of regulatory proteins can impair induction at low inducer concentrations. Inside the cell, enough repressor remains unaffected by inducer to repress most, if not all, the promoters on the plasmids controlled by LacI. Only high concentrations of IPTG can dislodge all the Lac repressor from its cognate binding sites on the DNA.

An excess of Lac Repressor is produced by plasmids containing a *lacI* gene, due to the increased dosage of the gene. In Table 10, the chromosomal expression of LacZ, also controlled by Lac Repressor, indicates the induction achieved by IPTG addition in the same way as globin gene expression, but is a much more sensitive and easily assayed indicator than globin expression. The LacZ expression, as a percent of that achieved by 1,000μM IPTG addition for SGE1662 versus SGE1464, demonstrates the induction of expression by low IPTG concentrations when lower LacI:plasmid ratios exist. In SGE1464, with an estimated ratio of LacI:plasmid of only 3:1, expression of chromosomal LacZ was induced at much lower concentrations of IPTG than was necessary to induce equivalent expression from SGE1662, which has an estimated ratio of LacI:plasmid of 10:1. Similar proportions of rHb1.1 are induced from SGE1464 at IPTG concentrations similar to that required for chromosomal LacZ induction (see above), indicating that this is an appropriate and accurate surrogate, since both are controlled by Lac Repressor in the same cell.

Table 10.

Percent Induction of chromosomal LacZ at various IPTG concentrations in the high copy number (1464) and medium copy number (1662) rHb1.1-producing strains.

µM IPTG	1662	1464
3	0.2	21.3
10	0.2	44
30	1.6	81.7
100	22.9	88.8
1.00	100	100

EXAMPLE 16

Characterization of chromosomal lacI alleles

DNA Isolation, PCR, Cloning, and Sequencing. Genomic DNA from strain C600 (*E. coli* Genetic Stock Center) was isolated from cells lysed by a freeze-thaw, lysozyme method. After SDS and proteinase K treatment, DNA was purified by phenol extraction and EtOH precipitation. The DNA was spooled from the EtOH solutions and rinsed. A map-based Polymerase Chain Reaction (PCR) cloning method was used to amplify the region upstream of the *lacI* gene. The Kohara map predicts that there is a PstI site at about 600 bp upstream from the *lacI* promoter. A 12 nucleotide-long oligo with a degenerate 5' half and the recognition site for Pst I at it's 3' end, and an oligonucleotide complementary to a region just downstream from the initiatation of translation site in the *lacI* gene (CBG23): AGT CAA GCT TAA CGT GGC TGG CCT GGT T were used in the PCR reaction. CBG23 contained the recognition site for HindIII. 50 pmole of CBG24 (5'-NNNNNNCTGCAG-3') and 10 pmole of TG45 (5'-CTGGCACCCAGTTGATCG-3') were used per 100 µl reaction. Pfu polymerase (Stratagene, La Jolla, CA) was used. 200 ng of purified C600 DNA was used. Reactions were performed in the Ericomp. The first cycle of the program consisted of 10 min. at 95°C, 5 min annealing at 50°C, 2 min extension at 72°C. 33 cycles consisting of 1 min at 95°C, 30 sec at 50°C, and 1 min at 72°C followed. The final cycle was 1 min at 95°C, 30 sec at 50°C, and 10 min at 72°C. Products were cleaned with the Magic PCR clean-up kit (Promega, Inc., Madison, WI) and digested overnight with PstI and HindIII. Digested DNA was cleaned again and ligated into pUC19 plasmid DNA that had been previously digested with *Pst*I and *Hind*III.

Ligation mixtures were transformed into JM109 (Promega, Inc.). Colonies containing inserts were identified and the the inserts were sequenced using the Sequenase, v.2.0 kit (United States Biochemical). Internal sequencing primers were used to complete the sequence of the inserts. A PCR primer containing a PstI site and homologous to the region 200 bp upstream from the *lacI* promoter was designed (CBG30): TAT ACT GCA GGG TAT TGG CTG TCT GAA T. Genomic DNA from RV308 (SGE2606) (Mauer et al. 1980. J. Molec. Biol. **139**:147-161 and Meyer et al. 1980. J. Molec. Biol. **139**:163-194), SGE1661, SGE1670, and SGE1675 was prepared as above. CBG23 and CBG30 were used to amplify the *lacI* promoter region. The PCR products were cloned as before into the *Pst*I and *Hind*III sites of pUC19 or pBC SK+ (Stratagene; La Jolla, CA). Insert-containing clones were identified by restriction digest and inserts were sequenced as before.

RNA Isolation and Primer-Extension Analysis.

RNA was isolated from bacterial cells by a hot-phenol extraction method. Cells were grown in shake flasks in M9 minimal media at 37°C. Primers were radiolabeled with γ^{32}P-ATP. Unincorporated label was not removed before annealing primer to RNA. RNA and labeled primer were mixed in annealing buffer and heated to 65°C for 10 minutes, followed by 50°C for 40 minutes. The mixture was then placed on ice and incubated for 15 minutes. Superscript II (Stratagene, Inc.) was used to synthesize cDNA according to manufacturer's instructions except for the following: 100 µg/ml vanadyl-ribonucleoside complex, 10 mM actinomycin D, RNasin were added. Reactions were precipitated with sodium acetate and ethanol. The precipitate was collected and washed with 70% ethanol and then dried. Pellets were resuspended in 3 µl water and 5 µl of sample buffer was added. Samples were heated at 65 °C for two minutes and then loaded on a 6% polyacrylamide sequencing gel adjacent to sequencing reactions of a plasmid with wild-type lacI that were performed using the same primer.

Cloning and Sequence of *LacI* Promoter Regions. *LacI* promoter region PCR products from strains SGE1661 an SGE1675 show two and sometimes three bands for most strains. RV308 has a large deletion of the *Lac* operon region but still has the upper band which was therefore believed to be an artefact of the PCR. Sequence analysis of this band revealed that it had a sequence unrelated to *LacI*. No sequence matches were found in database comparisons. The lowest bands are also believed to be artefacts. Two sizes of intermediate bands were observed. The lower was estimated to be 10 to 15 basepairs shorter than the upper. All strains which exhibited good control of expression had the lower band. An unanticipated result for PCR of SGE 1675, which had good control of expression, was the presence of

both bands indicating that at least a partial duplication of the *LacI* region may be present in this strain. PCR products were cloned into PUC19 or pBC SK+ at the *Hind*III and *Pst*I sites. Previously published *LacI* promoter region sequence data (Farabough, 1978, Nature 274:765-769 and Calos, 1978, Nature, 274-762-765) extends only 50 basepairs upstream of the start of transcription. Approximately 600 basepairs of additional sequence was obtained from C600. Smaller regions were cloned and sequenced from SGE1661, SGE1670, and SGE1675. The sequence of the $LacI^{Q1}$ allele is shown in Calos & Miller, 1981, supra. The upper, intermediate PCR band had a sequence similar to the wild-type with two minor, probably strain-related, differences. The lower band from SGE1670 and SGE1675 had a 15 basepair deletion. This creates a new -35 region which has the canononical sequence. This sequence is nearly identical to the published $LacI^{Q1}$ allele.

Kohara Map-Based Cloning Method for *Escherichia coli* DNA Regions.

A method utilizing Polymerase Chain Reaction (PCR) to amplify regions from purified DNA is described below. The Kohara *E. coli* restriction map (Kohara, *et al*, 1987) and the known sequence (Calos, 1978, and Farabaugh, 1978) of the *lacI* gene were used to design two oligonucleotide primers for the PCR.

Oligonucleotide 2 (CBG24) had the sequence 5'NNNNNNCTGCAG3' and should prime on any *Pst*I site in the genome. Oligonucleotide 1 (TG45) corresponded to known *lacI* sequence near the 5' end of the gene. In the first step, these two oligos were used to amplify a fragment of about 800 base pairs from *E. coli* strain C600 genomic DNA. Oligo 1 determines the specificity in this case. It is believed important to use a larger amount (50 pmole/100 µl reaction) of oligo 2 in order to compensate for the redundancy. 10 pmole/100 µl reaction of oligo 1 were used. This first PCR product was used in a second round of PCR which provides a second test of specificity. Oligo 2 plus a second internal *lacI* oligo, oligo 3 (CBG23) were used to amplify a subset of this region. Oligo 3 has a *Hind*III recognition site to allow cloning of this reaction product. PCR products were digested with *Pst*I and *Hind*III and cloned into pUC19 which was also digested with these enzymes. Clones with inserts were identified and sequencing was begun. Primers complementary to the vector and to known *lacI* sequence were used for the first reactions. Sequence determined from these reactions was used to design more oligonucleotides in order to complete the sequence.

Genomic DNA was then purified from SGE1661, SGE1670, and SGE1675. A primer that is 200 bp upstream from the *lacI* promoter was designed. A *Hind*III site was added to this oligonucleotide (CBG30). CBG23 and CBG30 were used to amplify the *LacI* promoter region from these three strains. PCR products were analysed by agarose gel electrophoresis where it was observed that the product from

SGE1670 had a small deletion relative to the product from SGE1661 (wild-type). In addition, two products, one of each size, were obtained from SGE1675. The PCR products were digested with *Pst*I and *Hind*III and cloned into pBC SK+ at the *Pst*I and *Hind*III sites. Candidates with inserts were identified and these were sequenced. It was determined that the smaller PCR product had nearly the same sequence as the *LacI*^{Q1} allele reported in 1981 by Calos and Miller. The larger product had the wild-type sequence.

References:

Calos, M.P. 1978. DNA sequence for a low-level promoter of the lac repressor and an "up" promoter mutation. Nature 274: 762-765.

Calos, M.P. and J.H. Miller. 1981. The DNA sequence change resulting from the I^{Q1} mutation, which greatly increases promoter strength. Mol. Gen. Genet. 183: 559-560.

Farabaugh, P.J. Sequence of the LacI gene. Nature 274: 765-769.

Kohara, Y., K. Akyama, and K. Isono. 1987. The physical map of the whole E. coli chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell 50: 495-508.

Measurement of *LacI* Transcript Level.

LacI transcript levels were assessed in JM109, SGE299, SGE1661, SGE1662, SGE1670, SGE1675. The results show that transcription levels in SGE299, SGE1670, and SGE1675 are similar to each other and are greater than in SGE1661, JM109 and SGE1662. The *LacI* transcript initiates at the same point in all strains. Quantitation of autoradiograms showed that the relative abundance of transcript of LacI in JM109, SGE299, SGE1662, SGE1670, SGE1675, respectively, was 1:5:7:17:17. JM109 contains *lacI*^{Q} on an F' factor which, is reported to result in ~100 copies of LacI per cell (Calos, M.P. ibid). SGE299 contains *lacI*^{Q1} on an F' factor, which since it has 5-fold more transcript than JM109, must therefore contain ~500 copies of LacI per cell. SGE1662 contains a lacI gene on the chromosome and on each of the estimated 100 copies of the expression plasmid, and would be estimated from the literature to have ~1,010 copies of LacI per cell. Since the transcript was 7-fold greater than measured from JM109, we would estimate ~700 copies of LacI per cell, in good accord with the expected ~1010 copies of LacI per cell. Both SGE1670 and SGE1675 contain lacIQ1 on the chromosome, and since they have 17-fold more transcript than JM109, must therefore contain ~1,700 copies of LacI per cell. Relative transcript

levels is one way to approximate the amount of regulator protein expected from mutations that increase promoter strength.

EXAMPLE 17

Effect of Exogenous Addition of Hemin on Soluble Hemoglobin Expression in SGE1483

The effect of addition of higher than normal concentrations of exogenous hemin on the soluble hemoglobin-like molecule, rHb0.0, accumulation in strain SGE1483 was examined by growing cells in 15 L fermentations for extended periods of time post-induction, with two different hemin feeding strategies. Soluble hemoglobin was measured using the IMAC procedure described above. Soluble hemoglobin was enhanced when additional hemin supplementation was included during extended incubation of the culture post addition of IPTG inducer (Table 11).

Induction of hemoglobin expression in 15L fermentations is normally supplemented by hemin additions at the time of induction, and at three and six hours post-induction. The volumes of hemin solution at each supplementation are 10, 13 and 17mls respectively, which result in a total hemin addition equal to 0.2g/L hemin in the fermentation tank. Additional hemin supplementation was achieved by a second and third 17ml addition at 9 and 12 hours post-induction during 16 hour induction incubations, resulting in a total of 0.37g/L hemin in the fermentation tank. The difference in soluble rHb0.0 levels at 10 hours post induction was not significant (Table 11), only after 16 hours post-induction did the additional hemin supplementation manifest a significant improvement in soluble rHb0.0 accumulation (Table 11). Hemin is hydrophobic, and known to precipitate in aqueous solutions. Therefore, the additional supplementation may help accumulate higher levels of rHb0.0 by increasing the total concentration, by maintaining a consistent soluble hemin level throughout the fermentation, or through other mechanisms. Supplementation strategies for hemin might also include , among the options, different bolus additions and timing of additions, continuous feeding, very high initial bolus additions, and alternative chelated hemin for improved solubility.

Table 11

Total g/L Hemin Added	g/L soluble rHb0.0 @ hours Post Induction	
	10	16
0.20	2.00±0.10	1.52±0.02
0.37	2.23±0.32	2.58±0.27

EXAMPLE 18

Effect of Exogenous Addition of Hemin on Soluble Hemoglobin Expression in SGE1464

The effect of addition of higher than normal concentrations of exogenous hemin on soluble hemoglobin-like protein, rHb1.1 accumulation in strain SGE1464 was examined by growing cells in 15 L fermentations for extended periods of time post-induction, with two different hemin feeding strategies. Soluble hemoglobin was measured using the IMAC procedure described above. Soluble hemoglobin was enhanced when additional hemin supplementation was included during extended incubation of the culture post-addition of IPTG inducer (Table 12).

Induction of hemoglobin expression in 15L fermentations is normally supplemented by hemin additions at the time of induction, and at three and six hours post induction. The volumes of hemin solution at each supplementation are 10, 13 and 17mls respectively, which result in a total hemin addition equal to 0.2g/L hemin in the fermentation tank. Additional hemin supplementation was achieved by a second and third 17ml addition at 9 and 12 hours post-induction during 16 hour induction incubations, resulting in a total of 0.37g/L hemin in the fermentation tank. The difference in soluble rHb1.1 levels at 10 hours post-induction was not significant (Table 12), only after 16 hours post-induction did the additional hemin supplementation manifest a significant improvement in soluble rHb1.1 accumulation (Table 12), as seen for rHb0.0 above.

Table 12

Total g/L Hemin Added	g/L soluble rHb1.1 @ hours Post Induction	
	10	16
0.20	1.14±0.21	1.53±0.28
0.37	1.28±0.23	1.95±0.26

All of the documents referred to herein are hereby incorporated by reference.

What is Claimed is:

1. A prokaryotic host cell for producing a heterologous polypeptide comprising:
a medium or high copy number plasmid comprising a regulatable expression unit encoding said polypeptide; and
a chromosomally-located gene encoding a regulatory protein capable of regulating said regulatable expression unit, expression of said regulatory protein being controlled by a strong promoter.

2. A prokaryotic host cell according to claim 1, wherein said plasmid is a high copy number plasmid.

3. A prokaryotic host cell according to claim 1, wherein said expression is repressible, said regulatory protein is a repressor capable of repressing said repressible expression unit, and expression of said repressor is controlled by a strong repressor promoter.

4. A host cell according to claim 1, selected from the group consisting of *Escherichia, Salmonella, Bacillus, Clostridium, Streptomyces, Staphylococcus, Neisseria, Lactobacillus, Shigella,* and *Mycoplasma.*

5. A host cell according to claim 4, wherein said host cell is *E. coli*

6. A host cell according to claim 1, wherein said regulatable expression unit is selected from the group consisting of lac, tac, trc, trp, ara, fru, gal and mal.

7. A host cell according to claim 6, wherein said regulatable expression unit is lac.

8. A host cell according to claim 5, wherein said expression unit is lac.

9. A host cell according to claim 3, wherein said repressor is selected from the group consisting of repressors of lac, tac, trc, trp, ara, fru, gal and mal.

10. A host cell according to claim 9, wherein said cell is *E. coli* and said repressor is LacI.

11. A host cell according to claim 10, wherein said strong repressor promoter is *lacI*^{Q1}.

12. A host cell according to claim 11, wherein the ratio of *lacI*^{Q1} to said plasmid is about 3:1.

13. A host cell according to claim 10, wherein said high copy number plasmid is selected from the group consisting of pSGE720, pSGE726 and pSGE728.

14. A host cell according to claim 3, wherein said host cell is *E. coli*, said repressible expression unit is lac-based, said repressor is LacI, and said strong repressor promoter is *lacI*^{Q1}.

15. A host cell according to claim 14, wherein said heterologous polypeptide is recombinant hemoglobin.

16. A host cell according to claim 15, wherein said recombinant hemoglobin is rHb1.1.

17. A method of producing a heterologous polypeptide comprising culturing, in the presence of an inducer, a bacterial host cell comprising:
a medium or high copy number plasmid comprising a regulatable expression unit encoding said polypeptide; and
a chromosomally-located gene encoding a regulatory protein capable of regulating said regulatable expression unit, expression of said regulatory protein being controlled by a strong promoter.

18. A method according to claim 17, wherein said plasmid is a high copy number plasmid.

19. A method according to claim 17, wherein said host cell is selected from the group consisting of *Escherichia, Salmonella, Bacillus, Clostridium , Streptomyces, Staphylococcus, Neisseria, Lactobacillus, Shigella,* and *Mycoplasma.*

20 A method according to claim 19, wherein said host cell is *E. coli.*

21. A method according to claim 20, wherein said regulatable expression unit is lac-based, said regulatory protein is LacI, and said strong promoter is *lacI*^{Q1}.

22. A method according to claim 17, wherein said regulatory protein is a repressor selected from the group consisting of repressors of lac, tac, trc, trp, ara, fru, gal and mal.

23. A method according to claim 17, wherein said regulatable expression unit is selected from the group consisting of lac, tac, trc, trp, ara, fru, gal and mal.

24. A method according to claim 17, wherein said heterologous polypeptide is recombinant hemoglobin.

25. A method according to claim 24, wherein said recombinant hemoglobin is rHb1.1.

26. A method according to claim 17, wherein said expression is repressible, said regulatory protein is a repressor capable of repressing said repressible expression unit, and expression of said repressor is controlled by a strong repressor promoter.

27. A method according to claim 17, wherein said host is cultured at a temperature of about 20°C to about 30°C.

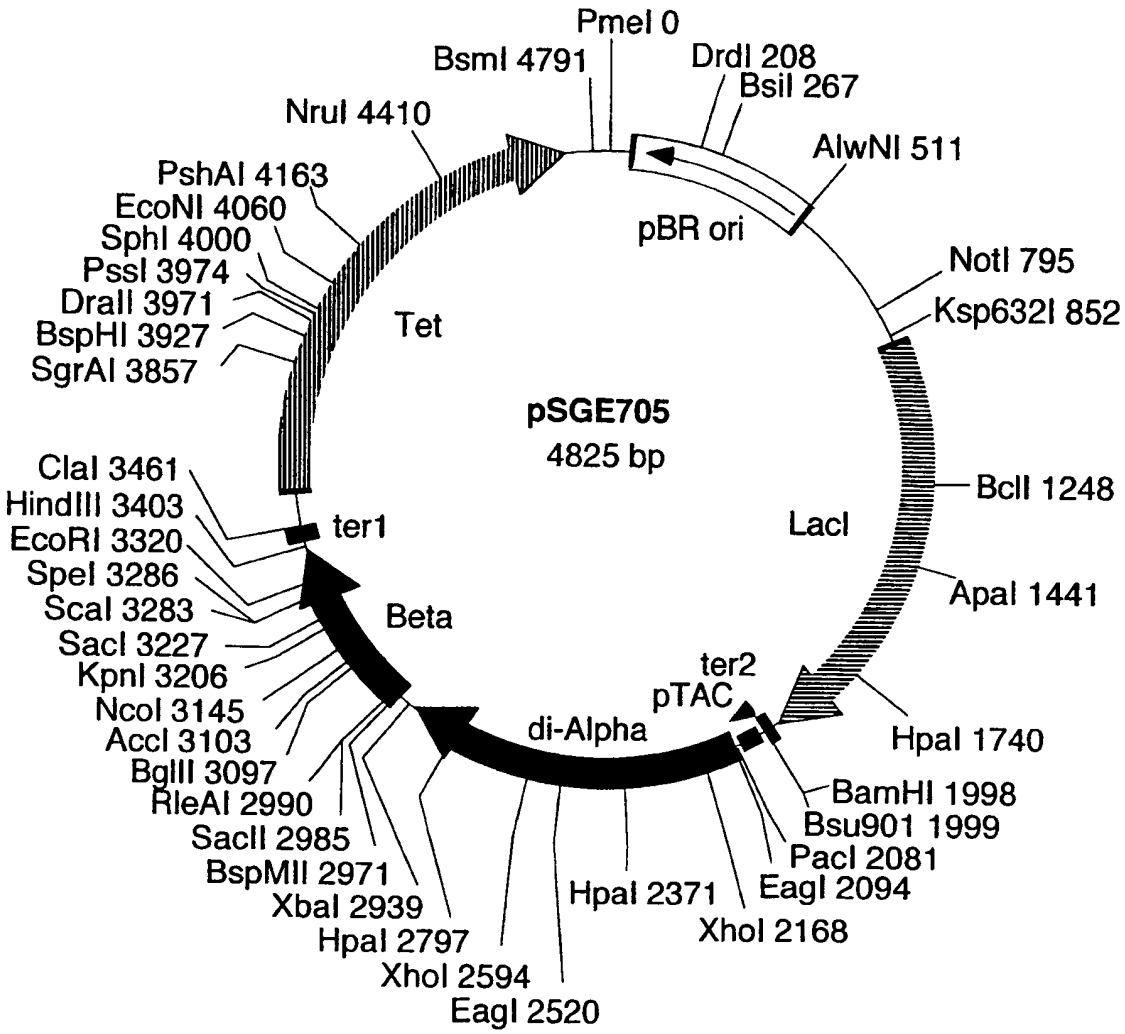
28. A method according to claim 27, wherein said temperature is about 24°C to about 28°C.

29. A method according to claim 28, wherein said temperature is about 26°C.

30. A high copy number plasmid selected from the group consisting of pSGE720, pSGE726 and pSGE728.

31. A plasmid according to claim 30, wherein said plasmid is pSGE720.

Fig. 1



RECTIFIED SHEET (RULE 91)
ISA/EP

Fig. 2a

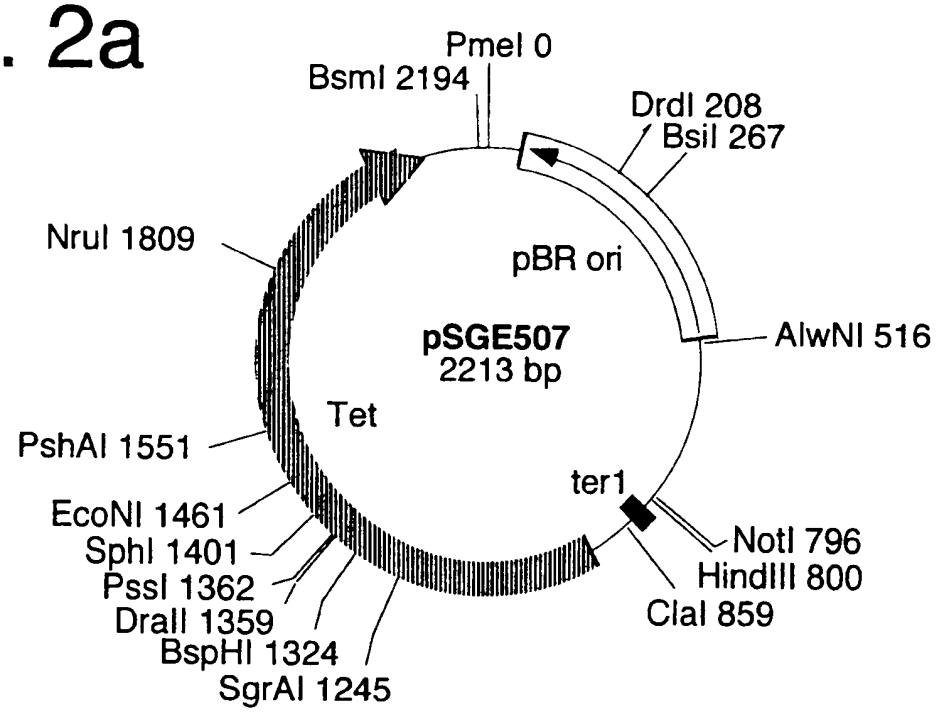
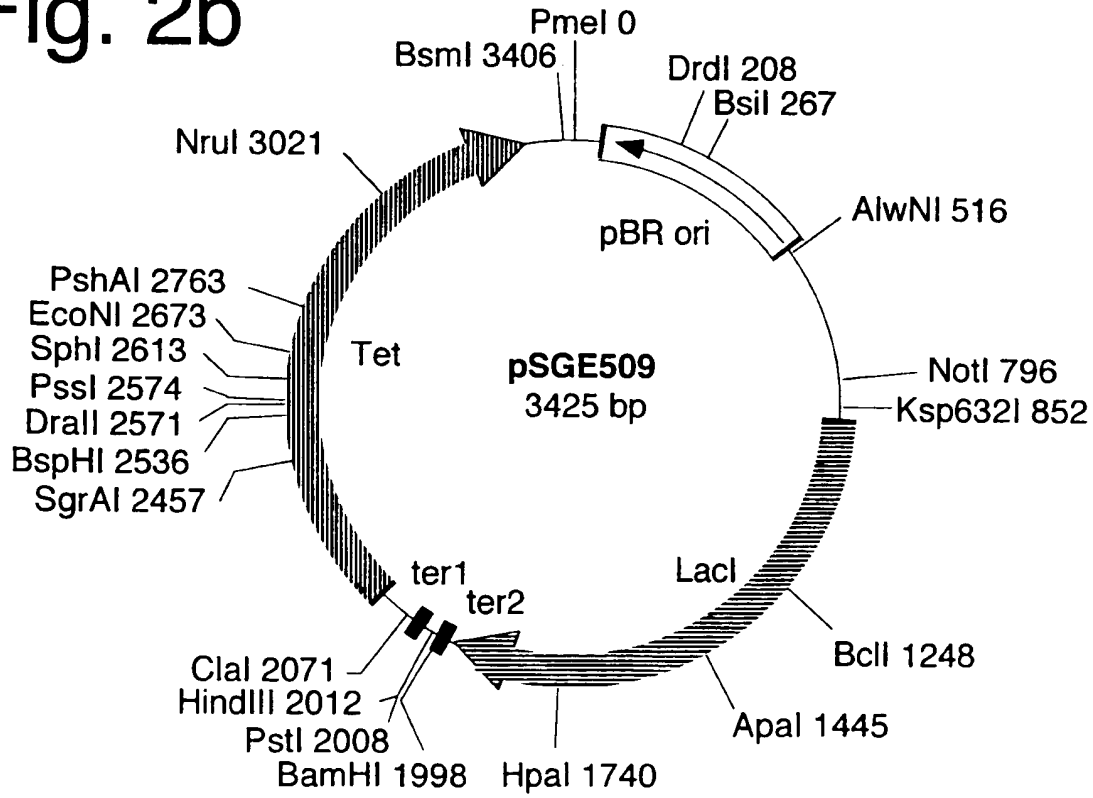
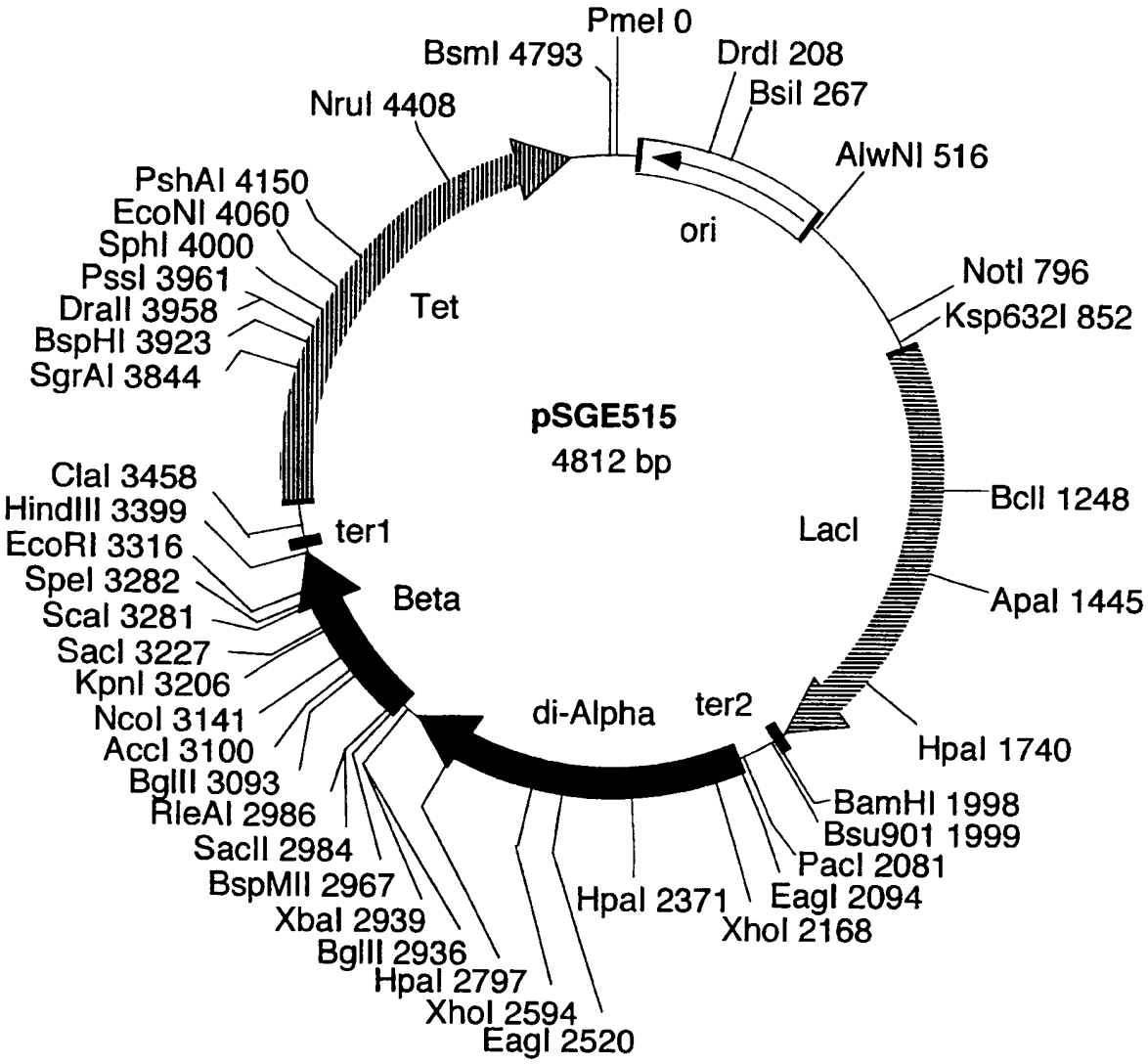


Fig. 2b



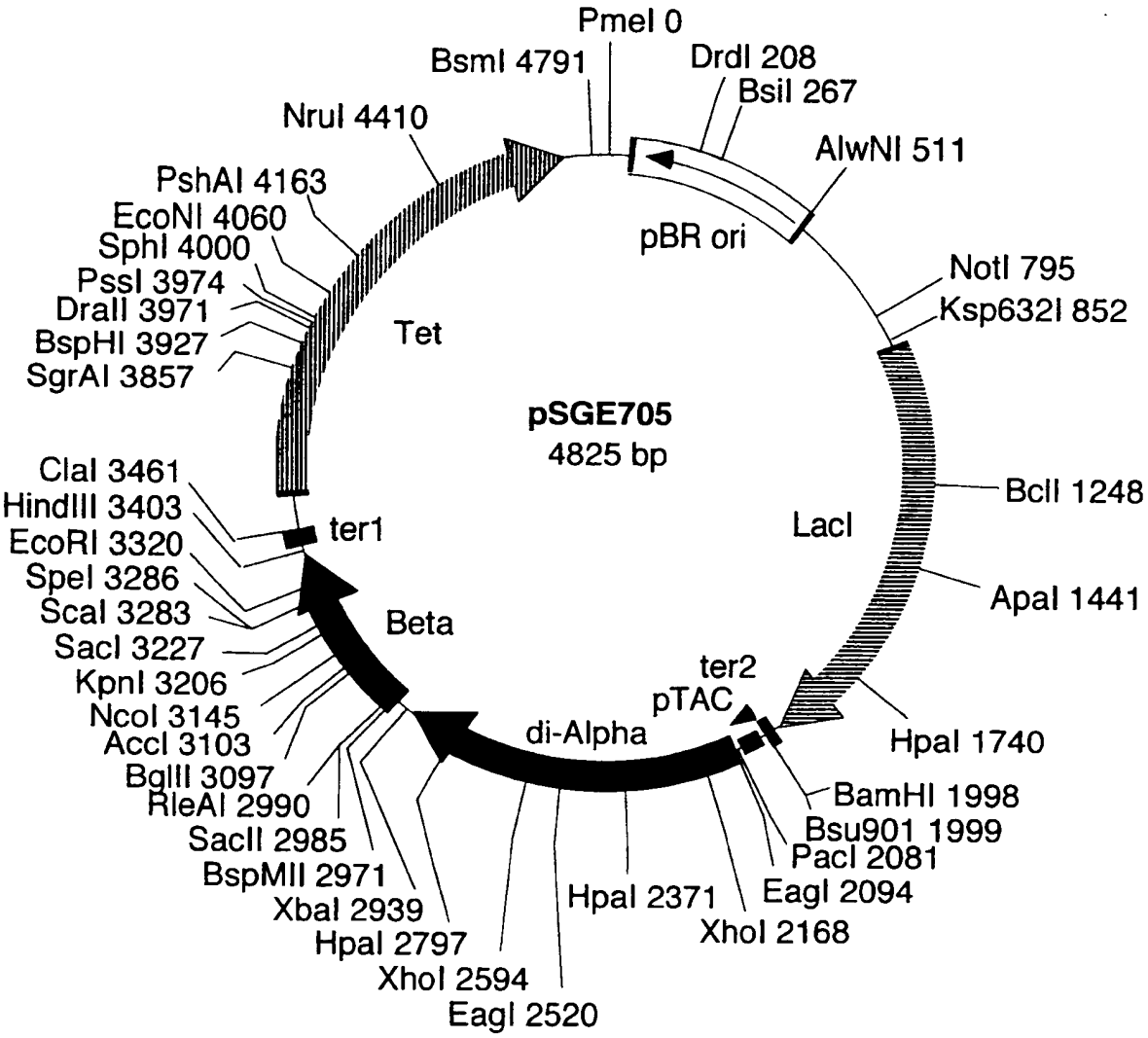
RECTIFIED SHEET (RULE 91)
ISA/EP

Fig. 2c



RECTIFIED SHEET (RULE 91)
ISA/EP

Fig. 2d



RECTIFIED SHEET (RULE 91)
ISA/EP

Fig. 3a

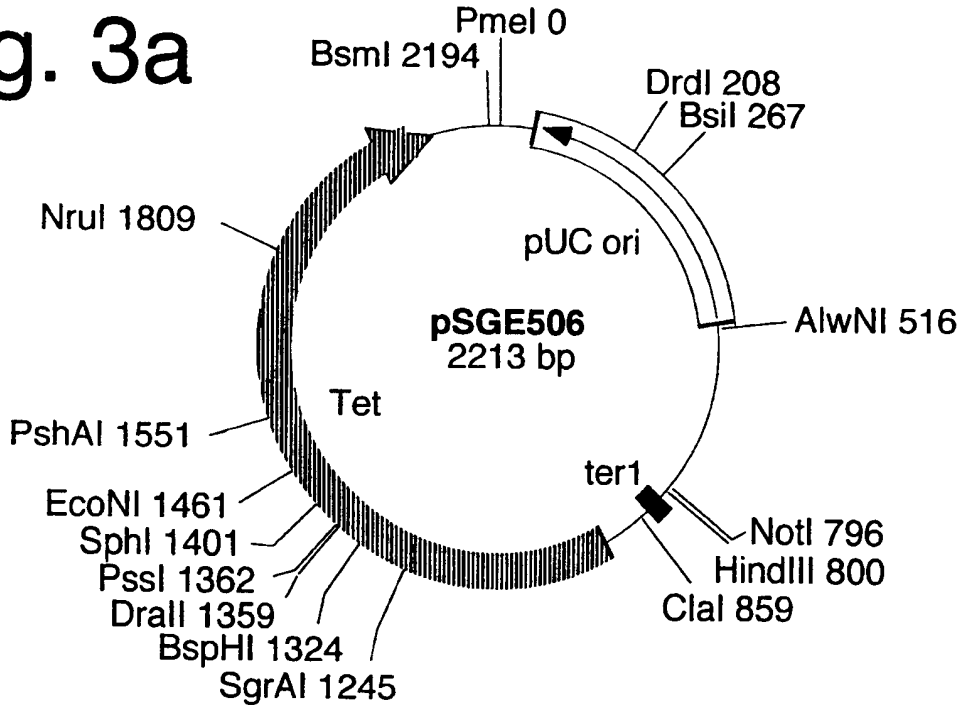
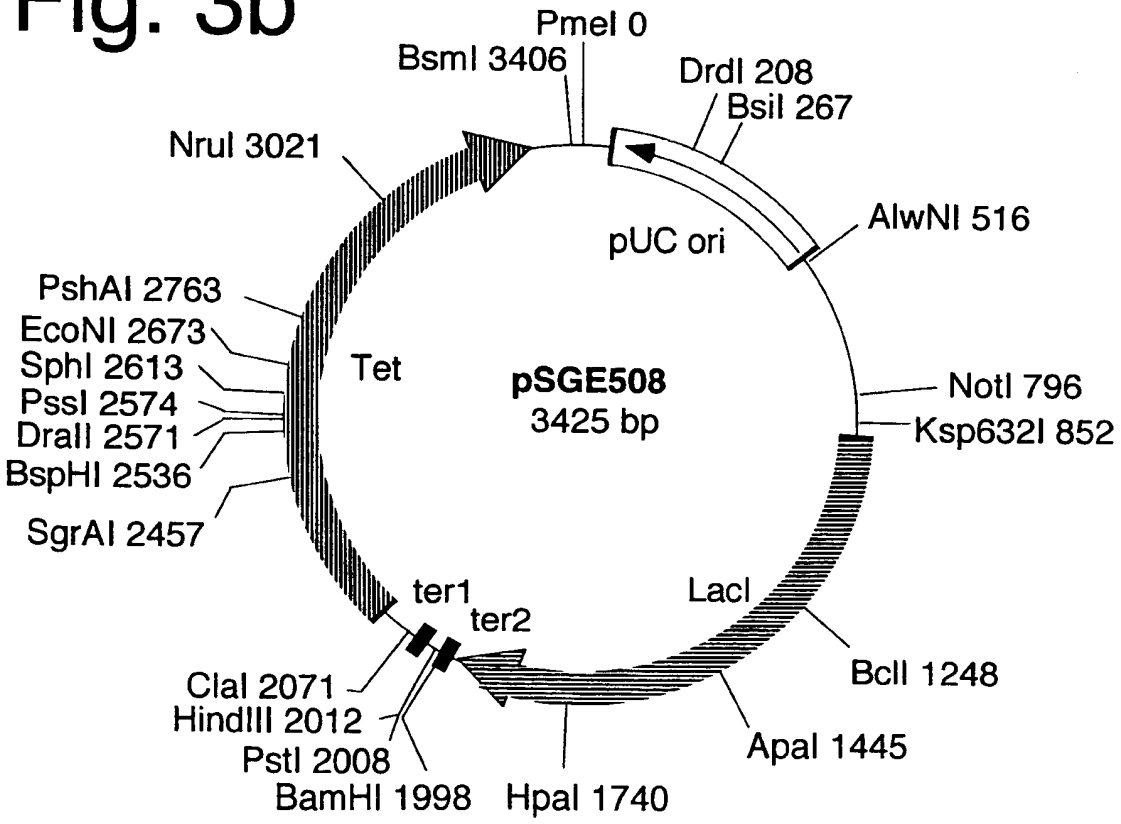
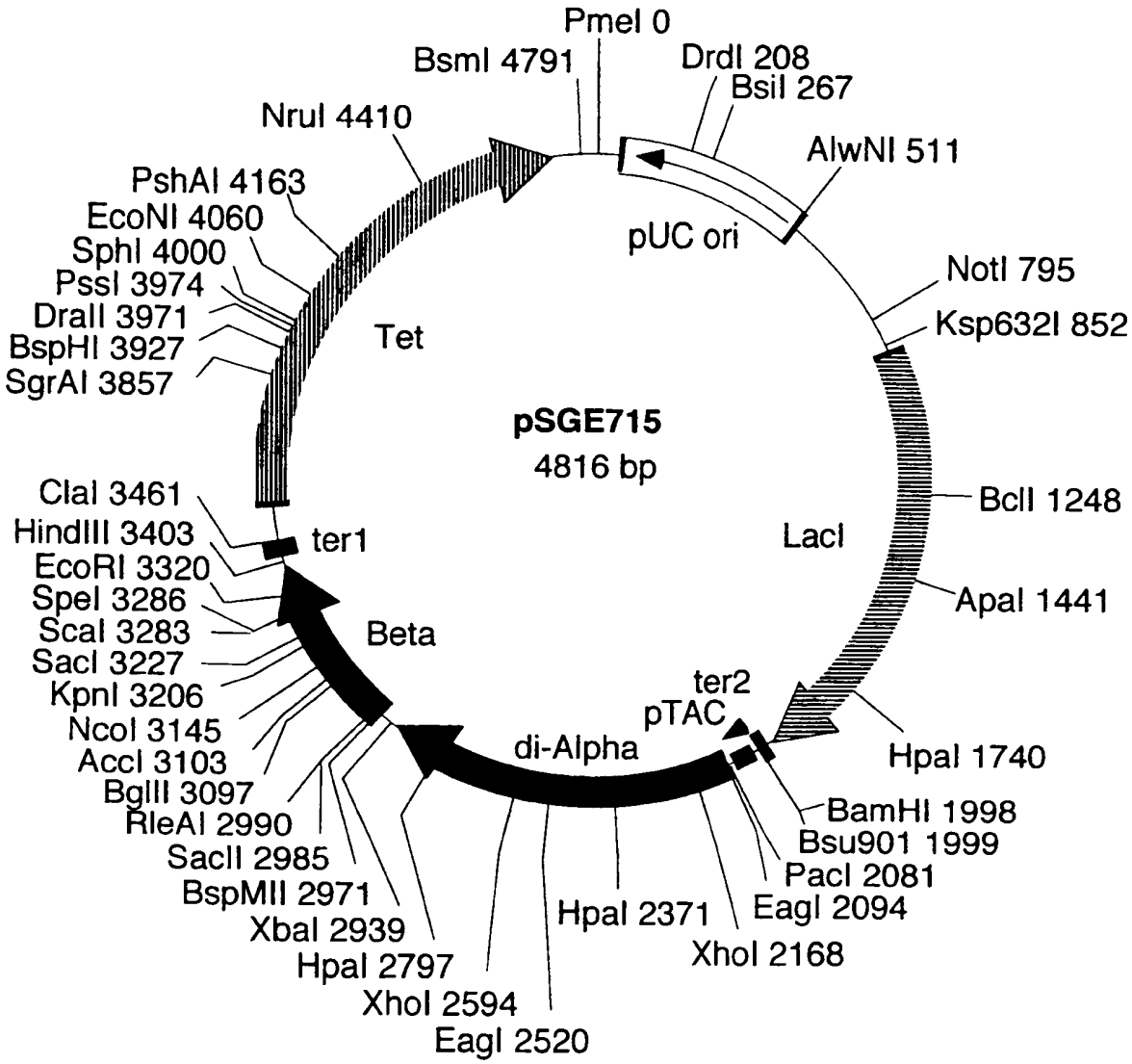


Fig. 3b



RECTIFIED SHEET (RULE 91)
ISA/EP

Fig. 3c



RECTIFIED SHEET (RULE 91)
ISA/EP

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 96/11600

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/63 C12N15/67 C12N15/70 C12N15/72 C12N15/12
C12N1/21 C07K14/805 //(C12N1/21,C12R1:19)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GENE, vol. 51, 1987, ELSEVIER SCIENCE PUBLISHERS,B.V.,AMSTERDAM,NL;, pages 255-267, XP002015969 M.R. STARK: "Multicopy expression vectors carrying the lac repressor gene for regulated high-level expression of genes in Escherichia coli"	1-11,14, 17, 19-23, 26-29
Y	see page 266, left-hand column, line 3 - line 6	2,12,13, 15,16, 18,24, 25,30,31
	see page 259, right-hand column, line 29 - page 264, right-hand column, line 47	

	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 October 1996	05. 11. 96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Hornig, H

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 96/11600

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GENE, vol. 25, 1983, ELSEVIER SCIENCE PUBLISHERS,B.V.,AMSTERDAM,NL;, pages 167-178, XP002015970 E. AMANN ET AL.: "Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in Escherichia coli" see page 176, left-hand column, line 25 - line 36	1-12,14, 17-23, 26-29

Y	MOLECULAR & GENERAL GENETICS, vol. 183, no. 3, 1981, SPRINGER INTERNATIONAL, AMSTERDAM, NL, pages 559-560, XP000605441 M.P. CALOS AND J.H. MILLER: "The DNA sequence change resulting from the IQ1 mutation, which greatly increases promoter strength" cited in the application see page 560, left-hand column, line 6 - line 11 see page 560, left-hand column, line 19 - line 20	1-12,14, 17-23, 26-29

X	GENE, vol. 69, 1988, ELSEVIER SCIENCE PUBLISHERS,B.V.,AMSTERDAM,NL;, pages 301-315, XP002015971 E. AMANN ET AL.: "Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli"	1,3-10, 17, 19-23, 26-29
Y	see the whole document	2,11-16, 18,21, 24,25, 30,31

X	EP,A,0 345 615 (BEHRINGWERKE AG) 13 December 1989	1,3-10, 17,19, 20,22, 23,26-29
Y	see page 4, line 53 - line 56 see page 9, line 1 - line 14	2,11-16, 18,21, 24,25, 30,31

	-/--	

4

INTERNATIONAL SEARCH REPORT

Interr nal Application No
PCT/US 96/11600

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PROC. NATL.ACAD SCI., vol. 80, 1983, NATL. ACAD SCI.,WASHINGTON,DC,US;, pages 21-25, XP002015972 H.A. DE BOER ET AL.: "The tac promoter: A functional hybrid derived from the trp and lac promoters" cited in the application	1,3-10, 17,19, 20,22, 23,26-29
Y	see page 21, left-hand column, line 47 - right-hand column, line 6 see page 23, left-hand column, line 3 - line 12; figure 3 see page 24, right-hand column, line 1 - page 25, left-hand column, line 15 ---	2,11-16, 18,21, 24,25, 30,31
X	EP,A,0 130 074 (GENENTECH INC) 2 January 1985 see page 16, line 1 - page 17, line 15 see page 21, line 22 - page 22, line 28 ---	1,3,4
Y	WO,A,90 13645 (SOMATOGENETICS INT) 15 November 1990 cited in the application see page 33, paragraph 3 - page 38, paragraph 3 ---	13,15, 16,24, 25,30,31
A	J. MOL. BIOL., vol. 219, no. 1, 1991, ACADEMIC PRESS,US, pages 45-59, XP000605448 J.W. DUBENDORFF AND F.W. STUDIER: "Controlling basal expression in an inducible T7 expression system by blocking T7 promoter with lac repressor" cited in the application see the whole document ---	1-31
A	PROC. NATL.ACAD SCI., vol. 59, no. 3, 15 March 1968, NATL. ACAD SCI.,WASHINGTON,DC,US;, pages 1259-1264, XP000605456 B. MÜLLER-HILL ET AL.: "Mutants that make more Lac repressor" cited in the application see the whole document --- -/--	1-31

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 96/11600

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PROC. NATL.ACAD SCI., vol. 85, December 1988, NATL. ACAD SCI.,WASHINGTON,DC,US;, pages 8973-8977, XP002015973 M. LANZER AND H. BUJARD: "Promoters largely determine the efficiency of repressor action" cited in the application see the whole document ---	1-31
A	NATURE, vol. 274, 24 August 1978, MACMILLAN JOURNALS LTD., LONDON,UK, pages 762-765, XP002015974 M.P. CALOS : "DNA sequence for a low-level promoter of the lac repressor gene and an 'up' promoter mutation" cited in the application see the whole document ---	1-31
A	J. BIOL. CHEM., vol. 267, no. 22, 5 August 1992, AM. SOC. BIOCHEM. MOL.BIOL.,INC.,BALTIMORE,US, pages 15869-15874, XP002015975 M.J. WEICKERT AND S. ADHYA: "A family of bacterial regulators homologous to Gal and Lac repressors" cited in the application see the whole document -----	1-31

Form PCT ISA 210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter onal Application No
PCT/US 96/11600

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0345615	13-12-89	DE-A-	3819463	14-12-89
		AU-B-	620873	27-02-92
		AU-A-	3608589	14-12-89
		JP-A-	2039891	08-02-90
EP-A-0130074	02-01-85	CA-A-	1264686	23-01-90
		JP-A-	60070085	20-04-85
		US-A-	4912046	27-03-90
WO-A-9013645	15-11-90	AT-T-	142692	15-09-96
		AU-B-	635744	01-04-93
		AU-A-	5675690	29-11-90
		CA-A-	2050601	11-11-90
		EP-A-	0402300	12-12-90
		EP-A-	0700997	13-03-96
		IL-A-	94361	31-07-95
		JP-T-	5503627	17-06-93
		US-A-	5545727	13-08-96